



**THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE  
FRANCHE-COMTE**

**PREPAREE A L'UNIVERSITE DE BOURGOGNE**

Ecole doctorale n°592

Ecole doctorale Lettres Communication Langues Arts (LECLA)

Doctorat de Sciences du langage

Par

Mme Angélica Leticia Cahuana Velasteguí

Expression de la sensorialité dans le discours spécialisé construit autour du chocolat en  
contexte équatorien

Thèse présentée et soutenue à Dijon, le 18 décembre 2023

Composition du Jury :

M. Dardo De Vecchi, Professeur associé Kedge Business School – Marseille, Rapporteur  
Mme Chrystelle Fortineau-Bremond, Professeure des Universités, Université Rennes 2, Rapporteur  
M. Laurent Gautier, Professeur des Universités, Université de Bourgogne, Co-directeur de thèse  
M. César Itier, Professeur des Universités, INALCO – Paris, Examineur  
Mme Anne Parizot, Professeure des Universités, Université de Franche-Comté, Examinatrice  
M. Stéphane Patin, Maître de Conférences HDR, Université Paris-Cité, Co-directeur de thèse  
Mme Mercé Pujol, Professeure des Universités, Université de Perpignan, Examinatrice

Volume 1



**Titre :** Expression de la sensorialité dans le discours spécialisé construit autour du chocolat en contexte équatorien

**Mots clés :** écologie du langage, linguistique de corpus, linguistique sensorielle, pragmatique, sémantique, analyse du discours, discours spécialisé.

**Résumé :** L'objectif de cette étude est de comprendre et comparer les manifestations de la sensorialité dans le discours spécialisé sur le cacao et le chocolat en contexte équatorien. Une description de l'écologie du langage a été présentée pour appréhender la relation entre le discours et la réalité des experts et des professionnels selon leur environnement physique, social et culturel. Afin d'étudier le processus sémantique des sensations et des émotions, trois sous-corpus ont été pris en considération : le sous-corpus ambiant (SCA) qui prend en compte l'expérience et le niveau d'expertise des universitaires et des professionnels dans le domaine du cacao et du chocolat ; le sous-corpus questionnaire (SCQ) qui provient de la production d'acteurs sociaux amateurs de chocolat dont l'intention de la communication est liée à l'expression corpus entrevue professionnelle (SCP) qui est produit par les professionnels étroitement liés à l'écosystème linguistique et aux activités du secteur, telles que la production et le marketing. Après la compilation des corpus, les résultats ont été analysés et interprétés à l'aide de la textométrie, de façon à modéliser les discours étudiés pour offrir un moule discursif représentatif des situations de communication étudiées en suivant une dimension heuristique de va-et-vient entre les sources empiriques et les résultats d'analyse textométrique connue sous le nom d'approche "bottom-up/top-down" (Brown et Yule, 1983 : 234-235). Finalement, ce travail offre une vue d'ensemble des différentes stratégies discursives utilisées pour exprimer la sensorialité, et donc pour évaluer l'empreinte du discours sensoriel dans les situations de communication

sensorielle à partir d'un stimulus ou d'une expérience évoquée ; et finalement le sous- respectifs situées dans leurs écosystèmes linguistiques

**Title :** Sensorial expression in the specialized discourse built around chocolate in an Ecuadorian context

**Keywords :** language ecology, corpus linguistics, sensory linguistics, pragmatics, semantics, discourse analysis, specialized discourse.

**Abstract :** The aim of this study is to understand and compare the manifestations of sensorial experience in the specialized discourse on cocoa and chocolate in the Ecuadorian context. A description of the language ecology was presented to understand the relationship between the discourse and the reality of experts and professionals according to their physical, social, and cultural environment. To study the semantic process of sensations and emotions, three sub-corpora were considered: the ambient sub-corpus (SCA) which considers the experience and level of expertise of academics and professionals in the field of cocoa and chocolate; the questionnaire sub-corpus (SCQ) which comes from the production of chocolate-loving social actors whose intention of communication is related to the sensory expression from a stimulus or an evoked experience; and finally, the sub-corpus of professional interviews (SCP) which is produced by professionals closely linked to the linguistic ecosystem and to the activities of the sector, such as production and marketing. After the compilation of the corpora, the results were analysed and interpreted using textometrics, in order to model the studied discourses to offer a representative discursive mould of the studied communication situations following a heuristic dimension of going back and forth between the empirical sources and the textometric analysis results known as the "bottom-up/top-down" approach (Brown et Yule, 1983: 234-235). Finally, this work offers an overview of the different discourse strategies used to express sensorial experience, and thus to assess the footprint of sensory discourse in communicative situations located in their respective linguistic ecosystems.



Je dédie ce manuscrit à

Olivier, Enrique, Flavie y Alejandro

# Remerciements

Le travail dont ce manuscrit est un des aboutissements représente un investissement qui n'aurait jamais pu porter ses fruits sans le soutien de plusieurs personnes que je tiens à remercier sincèrement.

Mes premières pensées vont à mon époux Olivier, qui a toujours été à mes côtés pour me motiver dans les moments de doute et à mes enfants Enrique, Flavie et Alejandro dont la patience a été mise à l'épreuve à d'innombrables occasions ; merci pour leur appui inconditionnel.

Je tiens aussi à exprimer toute ma reconnaissance à mon directeur de thèse, le Professeur Laurent Gautier et à mon co-directeur le Professeur Stéphane Patin, qui ont fait preuve de compréhension et d'empathie sans lesquelles le travail n'aurait pas pu être mené à terme.

Je voudrais également remercier les membres du Comité de suivi de thèse, pour leur confiance et les commentaires positifs partagés.

Pour finir, je remercie toutes les personnes qui, de près ou de loin, ont été impliquées dans le processus de production de ce travail et qui ont contribué à faire de cette thèse une réalité.

# Sommaire

Introduction	3
Première Partie : Cadre théorique et méthodologique pour la caractérisation des discours sensoriels autour du chocolat en contexte équatorien	11
Chapitre 1 : Écologie de la langue située dans la filière du chocolat en Équateur	13
Chapitre 2 : Le corpus en linguistique sensorielle	72
Deuxième Partie : Caractérisation analytique du corpus et modélisation	107
Chapitre 3 : Analyse textométrique et interprétation croisée des corpus considérés	109
Chapitre 4 : Empreinte du discours sensoriel dans les situations communicatives situées de l' <i>aficionado</i>	191
Conclusion	222
Bibliographie	239
Annexes	Volume 2



# Introduction

---

L'analyse du discours latino-américain est l'une des disciplines les plus innovantes en termes de conceptualisation de la langue et de sa relation avec le contexte social, politique et culturel de la région. Narvaja de Arnoux (2021) et Rodrigo-Mendizábal (2018) proposent un aperçu historique de l'évolution de l'analyse du discours latino-américain dans laquelle se retrouvent les différents courants, principalement européens, qui l'ont influencée. Cependant, aucun des deux auteurs fait référence directe au concept d'écosystème discursif qui a acquis une grande importance ces dernières années en raison de sa capacité à analyser la dynamique de la communication dans différentes sphères sociales. C'est ainsi que cette étude a la prétention de combler cette lacune en situant l'écosystème discursif au centre de son cadre théorique de manière qu'il devienne un sujet d'intérêt croissant dans le domaine de l'analyse du discours latino-américain. Ce concept découle de l'idée que tout système discursif est un système complexe qui comprend non seulement la langue utilisée, mais aussi les acteurs impliqués dans le discours, les contextes sociaux entourant le discours et les pratiques discursives qui se développent autour de lui.

La représentation d'un écosystème discursif trouve son origine dans le paradigme écologique de la biologie qui, au travers de cette métaphore, renvoie à l'idée que le discours évolue dans un environnement constitué de différents éléments, organismes et phénomènes qui coexistent et interagissent les uns avec les autres. La référence d'écosystème renforce l'idée d'une interaction étroite et vitale entre la langue et cet environnement social, culturel, économique et politique dans lequel le discours se construit. En ce sens, l'écosystème discursif est constitué d'un réseau de pratiques linguistiques, allant de l'utilisation de la langue de l'acteur social à la diversité des langues et des dialectes présents dans une communauté. L'interaction dynamique entre la langue et la construction du sens se situe dans le contexte social, culturel et historique

particulier à une situation de communication singulière. En d'autres termes, l'écosystème discursif englobe les diverses pratiques, discours et idéologies linguistiques qui façonnent la manière dont les gens utilisent la langue pour s'exprimer et communiquer entre eux au sein d'une communauté ou d'une société donnée.

Le concept d'écosystème discursif a été développé différemment selon les contextes linguistiques et culturels. Dans le monde anglophone, les premières études sur les écosystèmes discursifs sont apparues dans les années 1980 en réponse aux approches traditionnelles structuralistes qui se concentraient principalement sur la langue et la grammaire. Ainsi, des auteurs comme G. Kress (1985, 1988) ont décrit les discours comme des systèmes complexes allant au-delà de la phrase et du paragraphe, et ont proposé des concepts alternatifs proches de l'écosystème discursif pour analyser la complexité des systèmes de production langagière. Le paradigme d'écologie du discours, d'écologie du langage, d'écologie linguistique ou encore d'écilinguistique présenté par Haugen lors d'une conférence en 1970 et matérialisé dans une publication en 1972, s'est concentré sur l'analyse de la langue et de sa relation avec l'environnement physique et social dans lequel elle est mise en œuvre. Cette approche a permis aux chercheurs d'étudier les variations linguistiques dans différents contextes sociaux et géographiques, ainsi que les processus de changement, de production et d'interaction dans les dialectes et les langues. En ce sens, l'écologie linguistique se concentre sur l'analyse de la variation linguistique, de la diversité culturelle et de la relation entre la langue et la construction des identités et des pratiques sociales.

Dans le contexte hispanophone, plus spécifiquement en Amérique Latine, le concept d'écosystème discursif a été développé principalement à partir de l'analyse de la langue et de sa relation avec la diversité culturelle et linguistique des communautés latino-américaines, la notion d'écosystème discursif est proposée comme cadre théorique pour analyser la complexité des productions langagières dans les contextes sociaux, politiques et historiques de la région. Cette approche semble être particulièrement pertinente en raison de la complexité linguistique et culturelle des pays d'Amérique Latine où l'écologie linguistique est abordée à partir de différentes disciplines,

notamment la sociolinguistique, l'anthropologie, la linguistique et les études culturelles. L'importance de l'écologie linguistique dans ce contexte est due à la complexité linguistique et culturelle de la région, où des centaines de langues et de dialectes sont parlés et où la diversité linguistique est en constante évolution. Parmi les approches remarquables, il est possible de citer celles de l'analyse de la relation entre l'espagnol et l'anglais dans des contextes de bilinguisme à New York d'Ana Celia Zentella (1997) ; le travail de Luis Alejandro Ballesteros (2008) qui met en évidence la diversité du portugais brésilien selon les différents écosystèmes linguistiques dans lesquels il est produit ; et pour finir le concept d'écologie linguistique développé par Bastardas-Boada (2005, 2014, 2016, 2017) pour mettre en évidence la relation de pouvoir entre langue dominante et langue dominée. D'autre part, au cours de la dernière décennie, les études sur l'écosystème discursif ont connu un véritable essor dans la région, en particulier dans le contexte de la communication et des médias de masse. Il est aussi important de citer l'intérêt des linguistes allemands pour l'écologie linguistique d'Amérique Latine : Klaus Zimmermann (1999) et Yvonne Stork (2008) ont tous deux contemplé la diversité des langues hispano-américaines au travers du prisme de l'écologie linguistique pour défendre le besoin de protection de ces langues comme des espèces biologiques en danger d'extinction. Malgré la diversité thématique de ces études, dans leur grande majorité, elles ont pris en compte la relation entre l'écosystème discursif et la construction de l'identité culturelle des communautés étudiées afin de reconnaître l'importance de la langue dans la création d'imaginaires sociaux.

Par ailleurs, dans le contexte francophone, le concept d'écosystème discursif a été développé par des auteurs qui ont travaillé sur la relation entre langue, culture et société à travers l'analyse du discours politique et de son impact sur la construction de la réalité sociale. Claude Hagège (2006) relève l'importance de la diversité linguistique et culturelle pour le développement humain et social ; Louis-Jean Calvet (1999, 2019) souligne l'importance de la langue dans le processus de construction de l'identité individuelle et sociale, ainsi que l'utilité du concept d'écologie du langage comme cadre théorique de l'analyse du discours ; Jacques Guilhaumou (1996, 2000) analyse la

relation entre discours politique et construction de la réalité sociale en argumentant qu'elle ne peut être saisie que dans son contexte discursif ; Michel de Fornel en collaboration avec Jacqueline Léon (2000) aborde l'étude de la langue en relation aux pratiques culturelles au travers de l'ethnographie de la conversation ; Philippe Blanchet (2014) analyse la relation entre diversité culturelle et diversité linguistique au travers du plurilinguisme ; et pour finir cette liste non exhaustive, l'écologie (socio)linguistique de Jean Léo Léonard (2017) a le mérite de décrire la portée du paradigme de l'écologie des langues et d'en déterminer sa complexité. Malgré leur approche différente, ils soulignent tous l'importance de comprendre le contexte social et culturel dans lequel le discours est produit. Dans cette perspective, même s'il n'a pas toujours été nommé ainsi, l'écosystème discursif a été appréhendé comme un système composé d'acteurs, de rôles, de pratiques discursives et de contextes sociaux. Dans ces travaux, l'accent a été mis sur la relation entre le discours et la construction des imaginaires sociaux, en reconnaissant l'importance des médias, du cadre politique, de l'industrie culturelle et du système socio-économique où s'exprime le discours de l'opinion publique. Parallèlement à ce cadre général francophone, il convient de citer les travaux récents de l'équipe dijonnaise du laboratoire « Centre Interlangues Texte Image Langage » dans le cadre de laquelle a été menée cette thèse et qui considère l'écologie du langage comme un cadre théorique pertinent pour l'analyse du discours professionnel initialement appliqué au domaine viticole, mais rapidement diversifié vers des produits dont la filière professionnelle s'interroge sur le discours sensoriel produit par ses professionnels, ses experts et ses consommateurs.

Influencé par des facteurs politiques, culturels et économiques, le concept d'écosystème discursif n'est pas neutre. Par conséquent, les travaux se fondent actuellement sur la nécessité d'analyser l'écosystème discursif à partir d'une perspective critique, qui permet d'identifier les relations de pouvoir et les inégalités présentes dans la communication et la production de discours.

L'état d'avancement du concept d'écosystème discursif dans la sociolinguistique latino-américaine a été d'une grande importance au cours des dernières décennies, car il permet

de comprendre la relation entre la langue et le contexte socioculturel de la région. Bien qu'il se soit développé différemment selon les pays de la région, il constitue depuis les années 1980 un cadre théorique souvent utilisé par les chercheurs qui l'appliquent à différents domaines d'étude, tels que l'éducation, la politique, la littérature et la culture populaire.

Les concepts d'écosystème discursif et d'écologie du langage ont acquis une grande importance dans la sociolinguistique et l'analyse du discours latino-américaine en raison de leur capacité à comprendre la dynamique du discours et la relation entre la langue et le contexte socioculturel de la région. L'étude des écosystèmes discursifs est fondamentale pour comprendre comment le sens est construit et la réalité sociale produite dans notre ère mondialisée. Leurs caractéristiques sont différentes selon les contextes linguistiques et culturels, mais leur description permet d'analyser la variation linguistique, la construction de l'identité culturelle et des imaginaires sociaux, ainsi que les relations de pouvoir présentes dans la communication et la production de discours. Pour illustrer ces propos, il peut être considéré important d'analyser l'écosystème discursif d'un groupe social particulier afin de mieux comprendre ses valeurs, ses croyances et ses pratiques culturelles. Décrire l'écosystème discursif du lieu de travail moderne permet de mettre en évidence qu'il est façonné par un réseau complexe de discours liés à la productivité, à l'efficacité et à l'innovation. Un examen critique de l'écosystème discursif des médias peut révéler les façons dont le langage est utilisé pour construire et renforcer les stéréotypes sociaux et les inégalités.

Dans le cadre de cette étude, la construction de l'écosystème discursif du cacao et du chocolat en Équateur fait référence à l'interaction dynamique de la langue et du sens dans un contexte social, culturel et historique particulier qui remonte à l'époque coloniale, lorsque le cacao était déjà produit et exporté comme l'une des principales richesses de l'époque. La production et l'exportation de cacao ont commencé à contribuer de manière significative à l'économie équatorienne en 1780, lorsque l'Équateur faisait encore partie des colonies espagnoles et était connu sous le nom de

*Real Audiencia de Quito*<sup>1</sup>. Très tôt, un petit groupe de professionnels a constitué de véritables monopoles et accumulé un capital économique important. C'est à cette époque que l'on a commencé à appeler ces riches marchands *gran cacao* et que l'expression est devenue un dicton populaire en Équateur et d'autres pays d'Amérique Latine : *se cree el gran cacao - se ve a sí mismo como un gran cacao*, comme quelqu'un de meilleur que les autres. La fin du XIXe siècle marque le début de ce que l'on appelle le boom du cacao, les exportations passant de 372 433 quintaux en 1880 à 578 626 quintaux en 1899. En 1895, l'Équateur comptait 58,6 millions de cacaoyers. Dix ans plus tard, le pays en comptait 80 millions. Entre 1894 et 1913, l'Équateur est le premier exportateur mondial de cacao, avec 13 % de la production mondiale et 70 % des exportations du pays. Actuellement, l'Équateur s'est positionné sur le marché international comme le meilleur exportateur de cacao *fino de aroma* ; aussi bien la fève de cacao que le chocolat équatorien sont connus dans le monde entier (Abad, Acuña et Naranjo, 2020). L'intérêt de la population a considérablement augmenté, le cacao n'est plus uniquement cultivé pour être exporté, il est transformé en chocolat avant d'être exporté. L'écologie linguistique de la filière du cacao et du chocolat en Équateur s'est construite sur plus de 200 ans de culture et de traditions nationales indissociables d'un contexte socioculturel riche et conflictuel. La connaissance, même succincte, de ce patrimoine semble être un élément essentiel pour l'appréhension et la compréhension de l'écosystème linguistique dans lequel s'exprime la sensorialité dans le discours spécialisé construit autour du chocolat en contexte équatorien. Il est en effet possible d'observer empiriquement une différence marquée dans le discours des personnes qui sont en contact avec le cacao et le chocolat selon l'activité qu'ils exercent et le secteur de la filière dans lequel ils opèrent, en particulier dans le discours sensoriel. Différentes hypothèses pourraient être avancées sur ces disparités discursives. Elles pourraient être vérifiées selon des approches anthropologiques, sociologiques ou encore purement ethnographiques à partir de la variation des paramètres socioculturels de chaque locuteur. Elles pourraient aussi être étudiées dans le cadre d'une linguistique de corpus

---

<sup>1</sup> Comme convention pour l'ensemble de cet ouvrage, les expressions en espagnol et en anglais seront écrites en italiques sans traduction.

qui chercherait à illustrer le propos à partir d'exemples discursifs choisis. Cependant, le choix théorique et méthodologique de cette étude est de ne pas faire d'hypothèses anticipatoires sur la nature des disparités évoquées, mais de laisser au discours situé dans son contexte de production langagière le soin de révéler au chercheur les liens qui existent entre les stratégies discursives, les intentions de communication et les différents écosystèmes linguistiques des locuteurs. L'un des traits distinctifs de la méthode proposée est son approche ascendante : le processus d'analyse s'appuie principalement sur les textes représentatifs d'une situation de communication spécifique qui constituent le corpus compilé. Le chercheur découvre les stratégies discursives choisies par un profil déterminé de locuteurs selon les données quantitatives extraites du corpus à partir d'une approche corpus-driven (Glynn et Ficher, 2010 ; Biber, 2015 ; Adamou, 2016). La recherche est ainsi délibérément orientée vers une description spécifique du discours par le biais d'une analyse ascendante du corpus (bottom-up) basée sur une méthodologie inductive qui utilise les données statistiques auxquelles la textométrie permet d'avoir accès. Il convient de souligner qu'il s'agit de données quantitatives de fréquences d'usage d'unités linguistiques discrètes qui permettent au chercheur de modéliser les stratégies discursives mise en œuvre dans une situation de communication située et associée à un écosystème linguistique spécifique. Par conséquent, l'analyse s'apparente à une approche empirique qui permet la découverte progressive des stratégies discursives inscrites dans un contexte socioculturel spécifique afin d'avoir accès à l'intention de communication du locuteur représenté par le corpus compilé. Ainsi, à la fin de l'analyse, le chercheur peut proposer une collection de traits spécifiques décrivant le prototype représentatif du discours étudié qui, dans le cadre de cette étude, permet de proposer une modélisation des stratégies mise en œuvre pour partager une sensation, une émotion, qu'elle soit évoquée ou ressentie.

En conséquence, l'objectif principal de ce travail sera l'étude des stratégies discursives utilisées pour exprimer la sensorialité dans le discours produit par ceux qui sont en contact avec la filière du cacao et du chocolat : les agriculteurs qui cultivent le cacao, les professionnels qui transforment la matière première pour commercialiser un produit

fini, et les consommateurs. Le cadre théorique choisi pour cette étude est celui d'une linguistique cognitive sensorielle, formalisée et mise en œuvre sur d'autres produits par l'équipe dijonnaise, et qui interroge un corpus de textes situés compilés à partir de situations discursives professionnelles inscrites dans une écologie du langage propre à l'Équateur.

La description détaillée de ce cadre théorique, ainsi que l'écosystème linguistique auquel il s'applique, est proposée dans les deux premiers chapitres qui constituent la première partie de cette étude. Le premier chapitre débute par la conceptualisation de l'écologie du langage adoptée dans cette étude suivie d'une approche ethnographique de la spécialité et d'une explication succincte du processus de transformation de la fève de cacao en chocolat de façon à offrir au lecteur tous les détails qui constituent la scène équatorienne où se sont construits les discours étudiés. Le deuxième chapitre positionne l'analyse dans le cadre théorique précédemment décrit avant de présenter la méthodologie utilisée pour la compilation des textes recueillis dans les différents sous-corpus.

En tenant compte du discours étudié, l'analyse située (Condamines et Narcy Combes, 2015) réalisée pour connaître et interpréter les différentes stratégies d'expression du discours sensoriel utilisées dans les différents types de corpus et sous-corpus compilés permet de réaliser l'étude de l'utilisation de la terminologie ou d'expressions plus complexes en fonction de l'environnement, de l'activité à réaliser, et de l'intention de communication du locuteur. Le troisième chapitre présente l'analyse et la modélisation du discours des corpus ambiant et professionnel selon le type de production (cacao ou chocolat) et selon la langue utilisée (espagnol ou kichwa). Dans le cas du consommateur, appelé *aficionado*, une étude comparative sensorielle du discours avec et sans stimulus, c'est-à-dire avant et après la dégustation du produit, est décrite dans le quatrième chapitre afin de mettre en évidence les variations entre un discours produit à partir d'une évocation et un discours produit à partir d'un stimulus sensoriel.



L'organisation du travail présenté permet de connaître la situation de communication et l'intention de chacun des locuteurs, afin que l'étude des discours compilés révèle les variations que l'on peut observer dans les différentes situations de communication de la chaîne de production du chocolat, de la fève à la tablette en passant par la transformation et la vente.

L'approche analytique faite à partir des outils de textométrie propres à la linguistique de corpus offre une mise en évidence quantitative des différentes empreintes du discours sensoriel présentes dans ces situations communicatives. L'interprétation croisée des résultats de l'analyse textométrique situés dans les écosystèmes linguistiques respectifs offre une mise en évidence qualitative de ces différentes empreintes. En conclusion, l'analyse textométrique et l'interprétation croisée des corpus considérés aboutissent à trois contributions finales d'aspect théorique, méthodologique et pragmatique. Elles sont, présentées comme une possible valorisation des résultats.



# Première Partie :

## Cadre théorique et méthodologique pour la caractérisation des discours sensoriels autour du chocolat en contexte équatorien

---

Dans cette première partie d'une étude directement concernée par le discours sensoriel et situé dans un contexte équatorien autour du cacao et du chocolat, il convient de présenter la nature et la portée des variables sélectionnées. La description de l'écologie du langage, l'approche ethnographique de la représentation socioculturelle de la filière du cacao et du chocolat en Équateur ainsi que le processus de transformation rassemblés et discutés dans le premier chapitre sont des informations nécessaires pour permettre au lecteur d'accéder à l'essence même de ce que représente cette filière pour un Équatorien.

Ce prérequis guide le lecteur à travers la réflexion épistémologique proposée dans le deuxième chapitre où est exposé le positionnement théorique choisi pour justifier la mise en place des processus opératoires d'analyse textométrique et d'interprétation croisée des empreintes du discours sensoriel situées dans les écosystèmes linguistiques sélectionnés.

Le contenu présenté dans cette première partie est donc à la fois descriptif et opératoire. En effet, il décrit le cadre théorique dans lequel s'applique le mode opératoire spécifique :

Conceptualisation du paradigme de l'écologie du langage, suivie de la description de l'écosystème linguistique de la filière du cacao et du chocolat équatorien.

Conceptualisation d'une linguistique sensorielle située, suivie de la description du processus de compilation des corpus représentatifs des situations de communication professionnelles sélectionnées pour l'étude, ainsi que de leurs caractéristiques.

Introduction de la linguistique cognitive de corpus et de ses outils textométriques pour préparer la deuxième partie dédiée à l'analyse discursive et à l'interprétation croisée.

# Chapitre 1 : Écologie de la langue située dans la filière du chocolat en Équateur

---

Cette recherche s'inscrit dans le cadre de la linguistique sensorielle (Bach, 2021b) à partir de la construction d'une terminologie située (Condamines et Narcy-Combes, 2015) produite autour du cacao et du chocolat en Équateur, plus précisément dans la province de Pastaza située en Amazonie équatorienne. Dans notre cas, la compréhension de la sensorialité, à travers l'étude d'un discours situé, oblige une description suffisamment complète de la relation entre le cacao, le chocolat, l'homme et son environnement pour appréhender le contexte socioculturel qui détermine l'écologie actuelle de notre corpus et nous informe sur la représentation sociale des acteurs équatoriens impliqués dans le domaine spécialisé considéré.

Ainsi, la première section du chapitre présente l'écologie du langage comme cadre théorique général dans lequel s'inscrit l'approche ethnographique de la spécialité décrite dans la deuxième section. Il est important d'offrir au lecteur une information détaillée du processus de transformation artisanal et industriel du cacao, de la plante, des points clefs de son processus de transformation, de l'histoire du cacao en Équateur, et du chocolat actuellement produit dans le pays de façon à décrire le décor des situations de communication étudiées. Ces informations offriront au lecteur les connaissances nécessaires pour contextualiser les informations collectées de façon à pouvoir appréhender l'analyse et les interprétations de l'étude selon les différents écosystèmes dans lesquels le discours est construit.

## 1.1 Le paradigme de l'« Écologie du langage »

### 1.1.1 Définition, évolution et positionnement épistémologique de l'écologie du langage.

Actuellement, la recherche en sciences du langage s'est diversifiée dans de nombreuses approches poststructuralistes grâce, notamment, au développement des outils informatiques qui ont permis la naissance d'une linguistique outillée applicables à de nombreux champs interdisciplinaires et offrant, entre autres, de « nouveaux contextes de recherche pour les langues de spécialité » (Gautier 2019 : 371). La linguistique outillée met en œuvre l'outil informatique pour extraire des informations quantitatives d'un ensemble de textes représentatifs d'une situation de communication produite dans un environnement déterminé. Dans une approche « bottom-up/top-down » (Brown et Yule, 1983 : 234-235), ces informations quantitatives, issues de productions discursives authentiques permettent au chercheur d'avoir accès à des interprétations de la réalité (bottom up) qui sont ensuite validées et replacées dans l'environnement de production du discours analysé : l'écologie du langage (top down). Ce mouvement cognitif évite de poser des hypothèses *a priori* de façon à ce qu'elles apparaissent elles-mêmes de l'analyse du discours pour ensuite être vérifiées et validées dans leur écosystème sociolinguistique.

La perspective théorique articulant, depuis maintenant une cinquantaine d'années, les concepts associés à l'écologie du langage s'inscrit à la croisée interdisciplinaire des sciences du langage, sciences de l'information et la communication, et sciences cognitives au service d'une linguistique outillée, appliquée et située (Gautier et Bach, 2017). En effet, la considération et la description *in vivo* de l'écologie dans laquelle est mis en œuvre le langage authentique deviennent une étape importante du processus d'analyse et d'interprétation du discours donnant ainsi un cadre théorique général associé au concept de « langue-culture-milieu de spécialité » (Gautier 2019 : 374). Dans une société où l'interaction sociale met en œuvre des rapports de pouvoir, de statut,

d'idiosyncrasie ou encore d'hégémonie, il est difficile de réduire un échange à sa simple matérialité linguistique sans introduire l'écosystème social dans lequel il est produit. Cette réalité rend évident l'emprunt terminologique - écologie - fait à la biologie. Lechevrel, définit ainsi l'écologie comme une discipline scientifique qui fait partie de la biologie, mais surtout comme un rapport entre un être et son environnement physique, chimique et biologique.

L'écologie est une discipline de la biologie qui s'est constituée comme telle au cours du XIXe siècle... L'écologie du XXe siècle, que l'on nomme écologie moderne, a toujours pour but d'étudier les êtres vivants en rapport avec leur environnement physique, chimique et biologique. (Lechevrel, 2008 : 21-22)

Par conséquent, si l'écologie étudie la relation qui existe entre l'homme et son environnement, il est alors logique de considérer l'écologie du langage comme le rapport entre l'acteur social et son environnement physique, social et culturel. Depuis l'Antiquité, nous voyons comment le langage prend forme, se transforme, s'adapte aux besoins de l'environnement dans lequel il est produit et il évolue. Dans le cadre de cette évolution parallèle entre système linguistique, société et perception de la réalité, il est alors concevable de défendre l'approche théorique de l'écologie du langage comme l'un des héritages directs du relativisme linguistique. En effet, Leavitt conceptualise le relativisme linguistique comme le fait que chaque langue fait naturellement partie d'un tout socioculturel unique ; que la langue parlée reflète ou même détermine le mode de pensée de chaque individu, d'une manière distincte pour chaque communauté linguistique (2014 : 19). Et même si le débat reste virulent autour de l'hypothèse de Sapir-Whorf :

*there is a growing mass of research supporting some kind of influence of language specifics on conceptualization. Given that human beings are using specific languages, with all of the peculiarities of each, as tools to help think and communicate about the world and themselves thousands of times a day,*

*this is hardly surprising – it is simply a case of the tools used having some influence on the final product. (Leavitt, 2014 : 27)*

La relativité linguistique, parfois appelée hypothèse whorfienne, postule que les propriétés du langage affectent la structure et le contenu de la pensée et donc la façon dont les humains perçoivent la réalité. Une distinction est souvent faite entre les points de vue whorfiens forts, selon lesquels les catégories de pensée sont déterminées par la langue, et les points de vue faibles, qui soutiennent que la langue influence la pensée sans en déterminer entièrement la structure. (Danesi, 2021 : 10).

De Humboldt (Lafont et Peña, 1999) jusqu'aux récents travaux poststructuralistes en passant par Wittgenstein (1921), Sapir-Whorf (1956), Martinet (1960), Hymes (1974), Marcellesi et Gardin (1974), pour une sociolinguistique francophone ; sans être explicitement cité, de concept d'écologie du langage est en filigrane mettant en évidence qu'un acteur social interprète et perçoit sa réalité selon son environnement physique et socioculturel. L'un des premiers linguistes à avoir établi un lien entre la langue et l'environnement est Sapir, dans une conférence donnée en 1911 et publiée en 1912. Cependant, un des arguments whorfiens souvent cité pour illustrer la corrélation entre langage et perception de la réalité est la variété de mots utilisés en Inuktitut par les Inuits pour décrire la neige ou la glace, par exemple :

1. *qanik* : la neige qui tombe
2. *aputi* : la neige au sol
3. *pukak* : la neige cristalline au sol
4. *aniu* : la neige utilisée pour faire de l'eau
5. *siku* : la glace en général
6. *nilak* : eau potable glacée



7. *qinu* : neige fondue sur l'eau

Bien que cette théorie ait été fortement contestée par des auteurs comme Pullum (1991) qui ont accusé Whorf d'avoir mal interprété les différences linguistiques annotées par l'anthropologue Boas (1911 ; 1938), d'autres auteurs comme l'anthropologue Krupnik ont aussi soutenu l'idée que les langues des régions arctiques présentaient plus de mots pour décrire la neige que l'anglais (Krupnik, 2011 ; Krupnik et Müller-Wille, 2010). Il est aussi raisonnable de considérer que le fait d'avoir un seul mot en arabe pour la neige et la glace : *thalj*, n'est pas étranger à l'environnement dans lequel cette langue est parlée. Whorf considère qu'une société étroitement liée à ce genre d'environnement développe la langue qui rend compte de la réalité perçue. Ces descriptifs n'auront aucune, ou peu, d'équivalence directe dans un autre système langagier dont l'écologie n'offre pas des caractéristiques similaires aux environnements présents au nord du soixantième parallèle ou dans les contrées désertiques. Il conceptualise son hypothèse ainsi :

*We dissect nature along lines laid down by our native language [...] the world is presented in a kaleidoscopic flux of impressions which has to be organised by our own minds – and this means largely by the linguistic systems of our minds. We cut nature up, organise it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way – an agreement that holds throughout our speech community and is codified in the patterns of our language [...]. (Whorf, 1956 : 213)*

Il en est de même pour l'environnement socioculturel, Whorf a démontré que le système temporel des Hopis était différent d'un système linéaire continu à tel point que lorsqu'il demandait à un Hopi quels mots il utilisait pour désigner les jours passés (hier, avant-hier) ou les jours futurs (demain, après-demain), la réponse était qu'il ne comprenait pas la question. Comme le développe Villela-Petit dans son travail sur le temps dans la langue et la culture hopi (1992), les Hopis ne savaient pas ce que signifiait diviser le temps en jours passés et futurs. Leur conception cyclique et séquentielle du temps ne

leur permettait pas de comprendre la linéarité temporelle sous-entendue dans la question. Whorf précise ainsi que les êtres humains modulent leur système langagier selon la perception qu'ils ont de leur réalité, mais aussi que ce système conditionne d'une certaine manière leur façon de percevoir cette réalité. Par conséquent, deux acteurs sociaux différents peuvent manifester une vision différente de la même situation de communication dans laquelle ils produisent leurs discours, allant même jusqu'à considérer qu'une personne qui parle différentes langues, ou utilise différents registres de langue, peut appréhender la réalité de plusieurs façons. La diversité des différentes formes linguistiques reste très complexe, leurs variations ne peuvent seulement être comprises qu'à travers le prisme d'une linguistique structuraliste générativiste chomskienne. En effet, bien qu'elle soit l'un des courants les plus influents de l'histoire de la linguistique, elle présente certaines limites dans la compréhension de la variation linguistique. L'une des principales est qu'elle se concentre principalement sur la description de la compétence linguistique abstraite, c'est-à-dire la grammaire interne qu'un locuteur possède et qui lui permet de générer et de comprendre des phrases dans sa langue maternelle. Cependant, cette perspective ne tient pas compte de l'interaction du locuteur avec son environnement socioculturel et ses différents contextes d'utilisation. La variation linguistique est influencée par divers facteurs tels que l'âge, le sexe, la classe sociale, l'éducation et d'autres aspects culturels, historiques et politiques. Par exemple, certains contextes informels ou formels, des situations de contact linguistique ou des influences extralinguistiques peuvent contribuer à l'émergence de variations dans les usages linguistiques qui ne peuvent être expliquées dans la perspective de la linguistique structuraliste générativiste chomskyenne. En outre, il est important de noter que la variation linguistique ne se limite pas simplement à la sélection de différentes formes linguistiques parmi les options fournies par la grammaire d'une langue. La variation linguistique inclut également l'utilisation de différents registres, langues, sociolectes ou idiolectes, et peut s'accompagner de changements dans la prononciation, l'intonation et d'autres aspects qui ne sont pas abordés en profondeur dans la perspective générativiste chomskyenne. Même si la perspective générativiste chomskyenne fournit des outils utiles pour comprendre la

structure interne de la langue, elle n'offre pas une image complète de la variation linguistique dans sa complexité socioculturelle. La compréhension de cette variété de facteurs influençant la variation linguistique nécessite des approches complémentaires, telles que celles offertes par la sociolinguistique, l'anthropologie linguistique et d'autres approches interdisciplinaires incluant le concept d'écologie linguistique.

L'introduction de concepts comme la communauté linguistique que Martinet définit (1960) comme une communauté réunissant les individus qui parlent la même langue est une notion largement exploitée par Hymes (1974) pour soutenir ses critiques d'une approche essentiellement portée sur la fonction référentielle du langage et défendre une approche documentée, empirique et indissociable de la vie sociale pour appréhender la production langagière dans toute sa matérialité. Qu'elle soit identifiée comme *ethnographie de la communication*, *ethnolinguistique*, *sociologie du langage*, *anthropologie du langage* ou encore *sociolinguistique* selon les différentes sensibilités d'auteurs comme Marcellesi et Gardin en France (1974), cette approche semble un pilier théorique fondamental pour l'écologie du langage :

*In a nutshell, the vitality of languages depends on how regularly members of the populations associated with them get to use them in various settings. The consequences vary depending on whether they are used in all domains of the speakers' knowledge or experiences, only in some but not in others, or in none at all. Thus, the vitality of languages depends on what their speakers (do not) do with them. Therefore, there can be no accurate account of the ecology of a language that does not factor in the population associated with it and the communicative settings in which it evolves, including the particular socioeconomic structure that leads speakers to decide, where more than one language (variety) is involved, which one is more advantageous to them [...].*  
(Mufwene, 2018 :74)

Mufwene manifeste que la vitalité et le dynamisme des langues dépendent aussi de l'usage que les interlocuteurs en font, de ce que l'acteur social décide d'en faire ou de

ne pas en faire. Il considère que pour expliquer l'écologie des langues, il est nécessaire de prendre aussi bien en compte les caractéristiques de la population que l'environnement communicatif dans lequel elle se développe en mettant l'accent sur les aspects socio-économiques qui interviennent dans le choix des stratégies discursives de l'acteur social. Les acteurs sociaux doivent être capables d'utiliser leurs systèmes langagiers à bon escient afin de communiquer efficacement et de transmettre le message souhaité selon leur intention de communication.

L'écologie des langues, également appelée écologie linguistique, écologie du langage ou écolinguistique, dresse un cadre théorique général pour l'étude de ces relations entre les langues et les facteurs environnementaux aussi bien physiques que sociaux. « Différentes acceptions sont donc observables et l'écolinguistique offre de très nombreux travaux » (Lechevrel, 2008 : 17). Un rapide état des lieux permet de présenter comment divers chercheurs ont interprété l'écologie du langage.

Le courant est en plein essor aux États-Unis (S. Mufwene à Chicago, les chercheurs du B. L. Center), au Canada (W. Mackey), en Australie (P. Mühlhäusler) et dans certains pays européens comme les pays germanophones (A. Fill, J. Døør et J. Bang) ou la péninsule ibérique (A. Bastardas-Boada). En France, de nombreux chercheurs s'y intéressent également, mais seul L.-J. Calvet lui a consacré un ouvrage (Calvet, 1999). (Lechevrel, 2008 : 17)

Il existe aussi différents centres de recherches écolinguistiques dans le monde :

Bielefeld (Allemagne), Graz (Autriche), Odense (Danemark), Adélaïde (Australie) et Chicago (U.S.A.), et pour les acteurs : J. Chr. Bang et J. Døør (Odense), R. Alexander (Vienne), A. Fill (Graz), Finke and W. Trampe (Bielefeld), M. A. K. Halliday (Sydney), Mühlhäusler (Adélaïde) et A. Makkai (Chicago). (Lechevrel, 2008 : 17)

Comme précédemment commenté, il existe un lien étroit entre relativisme linguistique et écologie du langage. Ce sont deux approches qui se concentrent toutes deux sur

l'interrelation entre la langue et le contexte social et culturel dans lequel elle est utilisée. Le relativisme linguistique repose sur l'idée que la langue et la culture sont interconnectées de telle sorte que le sens et la signification des mots et des expressions sont relatifs à la culture dans laquelle ils sont utilisés. Cela signifie qu'il n'existe pas de conception universelle de ce qui est "bon" ou "mauvais" dans une langue, mais que les normes sont basées sur la culture dans laquelle elles sont utilisées et interprétées. L'écologie du langage, quant à elle, se concentre sur l'influence des facteurs environnementaux, sociaux et culturels sur l'évolution et l'utilisation des langues. L'écologie linguistique considère la langue comme un système adaptatif qui réagit aux changements de l'environnement et de la société dans lesquels elle est utilisée, et ne peut être comprise en dehors de ce contexte. Les deux approches soulignent l'importance du contexte socioculturel dans la compréhension de la langue et donc de la réalité dans laquelle elle est utilisée. De même, elles reconnaissent les processus dynamiques et changeants de la langue, la parole étant un phénomène flexible, adaptable aux contextes dans lesquels elle opère. En bref, le lien entre le relativisme linguistique et l'écologie des langues réside dans la reconnaissance de la relation étroite entre la langue et le contexte dans lequel elle est utilisée, et dans la nécessité de comprendre la langue comme une construction sociale et culturelle, plutôt que comme un objet isolé et décontextualisé. Sapir considère ainsi, plus précisément, que la langue est étroitement liée à l'expérience, qu'elle représente bien plus qu'un simple code et qu'elle nous permet de comprendre et d'interpréter la vision du monde de ses locuteurs.

C'est le vocabulaire d'une langue qui reflète avec le plus de netteté le physique et social de ses usagers. Nous pouvons en fait définir le vocabulaire total d'une langue comme un inventaire complexe de toutes les idées, intérêts et occupations qui retiennent l'attention d'une communauté ; et si nous disposions, pour une tribu donnée, d'un dictionnaire complet de sa langue, nous pourrions presque entièrement déduire le caractère du milieu physique et les traits marquants de la culture des usagers. (Sapir, 1968 : 75)

Haugen (1972), considéré comme le père fondateur de l'écologie linguistique, articule sa recherche autour de paramètres écologiques, qui se fondent sur l'interrelation qui existe entre une ou plusieurs langues et leur environnement linguistique, culturel, social, politique, géographique et naturel, entre autres.

*Language ecology may be defined as the study of interactions between any given language and its environment [...]. Language exists only in the minds of its users, and it only functions in relating these users to one another and to nature, i.e. their social and natural environment. Part of its ecology is therefore psychological: its interaction with other languages in the minds of bi- and multilingual speakers. Another part of its ecology is sociological: its interaction with the society in which it functions as a medium of communication. The ecology of a language is determined primarily by the people who learn it, use it, and transmit it to others. (Haugen, 1972 : 325)*

Halliday, dans sa conception du langage, se rapproche de l'hypothèse de Sapir-Whorf, et affirme que le langage construit la réalité :

*The categories and concepts of our material existence are not 'given' to us prior to their expression in language. Rather, they are construed by language, at the intersection of the material with the symbolic. Grammar, in the sense of the syntax and vocabulary of a natural language, is thus a theory of human experience. It is also a principle of social action. In both these functions, or metafunctions, grammar creates the potential within which we act and enact our cultural being. This potential is at once both enabling and constraining: that is, grammar makes meaning possible and also sets limits on what can be meant. (2006 :179)*

Halliday nous parle de l'organisation fonctionnelle interne du système linguistique, il soutient l'idée que les constructions idéologiques sont fortes et construites par notre syntaxe et notre connaissance terminologique, il considère que l'être humain situé dans son environnement social est l'élément central de la structure du langage.

Cette association écolinguistique du langage humain et de son environnement présente un large champ de définitions à travers différents travaux tels que ceux qui parlent de la diversité linguistique du point de vue de Haugen. En d'autres termes, ces travaux soutiennent l'idée de Haugen que les langues, comme les différentes espèces de plantes et d'animaux, sont en état d'équilibre, qu'elles se complètent, qu'elles se confrontent et que leur existence dépend les unes des autres. Le langage devient l'élément nécessaire à l'existence de cette relation entre l'homme, la société et la nature. Le fonctionnement et le développement du langage contribuent à la formation d'un écosystème dans un environnement qui n'est autre que la réalité sociale de l'individu.

Selon le système écologique de Haarmann, il y a sept variables environnementales qui déterminent le comportement linguistique, ce sont les variables ethnodémographique, ethnosociale, ethnopolitique, ethnoculturelle, ethnopsychologique, interactionnelle et ethnolinguistique (1990 : 109-110). Il affirme que ces variables sont interdépendantes et forment un système écologique constituant un tout, tout en se démarquant d'une analogie directe avec la biologie par l'essence même des variables considérées qui sont fondées sur des concepts psychologiques, sociologiques, culturels, politiques et linguistiques difficilement comparables aux variables des biologistes contrairement à l'approche de Mufwene. En effet, Mufwene (2001) se concentre de son côté sur l'évolution des langues et du langage en considérant un positionnement épistémologique qui prend en compte l'évolution darwinienne comme fondement théorique de son évolution. Par la suite, ses études ont fait place à d'autres recherches pour analyser le langage dans le cadre théorique de l'écologie du langage de Haugen qui s'appuient sur deux aspects généraux sous le titre de l'écologie du langage : un système de classification des langues proche de la taxonomie, mais aussi un concept généraliste prônant un modèle linguistique inclusif pour synthétiser l'ensemble des caractéristiques des différentes langues.

Ce concept d'écologie du langage est aussi partagé par Calvet (1999) qui affirme que le langage n'est pas un objet qui peut être isolé de son contexte de production, que la communication orale ne se produit pas simplement au moyen de séquences de sons,

mais qu'elle met en œuvre une pratique sociale, la production langagière, au sein de la vie sociale où se mélangent de nombreuses pratiques sociales indissociables de leur environnement. Il développe son idée en insistant sur le caractère évolutif et adaptatif du langage réagissant aux changements de son environnement, et adoptant, comme Mufwene, une vision darwiniste qui consiste à considérer qu'un écosystème biologique évolue et s'adapte aux variations de son environnement :

Les pratiques qui constituent les langues d'une part, et leur environnement d'autre part, forment un système écolinguistique, dans lequel les langues se multiplient, s'entremêlent, varient, s'influencent, se concurrencent ou convergent. Ce système est en relation avec l'environnement. À tout moment, la langue est soumise à des stimuli externes auxquels elle s'adapte. La régulation, que je définirai comme la réaction à un stimulus externe par un changement interne qui tend à neutraliser ses effets, est donc une réponse à l'environnement. Cette réponse est avant tout la simple addition de réponses individuelles, de variantes qui, au fil du temps, conduisent à la sélection de certaines formes, de certaines caractéristiques. En d'autres termes, il existe une action sélective de l'environnement sur l'évolution du langage. (Calvet, 1999 : 35)

La notion de régulation est ici primordiale. Calvet la présente comme une réponse à l'addition de réponses individuelles, au pluriel, apportant un argument irréfutable pour l'utilisation nécessaire d'une linguistique outillée qui permet la quantification objective de réponses individuelles comme évidence d'une sélection naturelle. Par analogie au principe d'évolution darwinienne, la sélection de certaines formes dépend de leurs occurrences d'usage qui peuvent être déterminées dans un corpus représentatif d'une situation de communication située comme celle du langage sensoriel associé à un produit de consommation : le chocolat.

Les études de Bastardas-Boada adoptent une vision systémique de l'organisation du monde tout en considérant l'écosystème comme un facteur déterminant dans la production langagière : « *human beings are units that both form and live in complex socio-political ecosystems, which in turn have an influence on their language*



*behaviours and can determine specific evolutions depending on the intervening factors in each case* » (2017 : 30). Son approche écolinguistique repose sur la notion d'écosystème, l'approche circonstancielle, la dynamique de la pratique linguistique, la coévolution et la continuité. Selon cet auteur, l'écologie permet de décrire, d'analyser et de comprendre la réalité de l'évolution de la diversité linguistique à partir de l'écologie des langages en contact fondée sur une approche psychologique, sociologique et politique en accord avec la conceptualisation de Haugen : l'écologie de la langue est tout aussi intimement liée à l'environnement qu'au socioculturel.

*The true environment of a language is the society that uses it as one of its codes. Language exists only in the minds of its users, and it only functions in relating these users to one another and to nature, i.e. their social and natural environment. Part of its ecology is therefore psychological: its interaction with other languages in the minds of bi- and multilingual speakers. Another part of its ecology is sociological: its interaction with the society in which it functions as a medium of communication. (2001 :57)*

Haugen précise ici l'importance d'une approche psychologique et sociologique de l'écologie du langage de façon à rendre compte des interactions entre différents langages et entre langage et société, lieu de rencontre des variations sociolinguistiques. Cet espace bidimensionnel nous permet d'apprécier les rapports de force entre une langue majoritaire véhiculant une culture dominante et une langue minoritaire représentative d'une culture dominée, mais aussi l'impact d'un bilinguisme bien marqué sur la production langagière de l'acteur social dans l'espace-temps d'une situation de communication spécifique. Dans l'écologie du langage équatorienne, cet espace bidimensionnel rend compte des rapports de forces entre l'espagnol, la langue majoritaire, et les différentes langues ancestrales définies comme minoritaires.

En outre, l'écologie des langues nous permet de prendre en considération l'évolution naturelle des langues et de la diversité culturelle en relation directe avec l'histoire de l'Équateur. Au cours de la période de colonisation, les langues européennes se sont

imposées aux langues indigènes dans tout le continent amérindien. La conquête de l'Amérique s'est accompagnée d'une conquête culturelle et linguistique. En Équateur, la langue de colonisation - l'espagnol - s'est imposée en reléguant et réduisant les langues autochtones à une utilisation locale identitaire, comme le consacre l'article deux de la constitution équatorienne<sup>2</sup> :

*El castellano es el idioma oficial del Ecuador, el kichwa y el shuar son idiomas oficiales de relación intercultural. Los demás idiomas ancestrales son de uso oficial para los pueblos indígenas en las zonas donde habitan y en los términos que fija la ley*<sup>3</sup>. (Art. 2 Constitution de la République d'Équateur, 2008 : 8)

Il est nécessaire de souligner qu'au moment de l'indépendance, il n'y a eu aucun changement à cet égard. Bien au contraire, la dévalorisation des langues indigènes a été renforcée par les gouvernements créoles successifs, entraînant l'affaiblissement et la disparition de nombreuses langues-cultures autochtones, de modèle de pensées et de visions différentes du monde. Même si en 2008 un effort a été fait au niveau de la constitution équatorienne, la *Figure 1* : Territoires des peuples et nationalités autochtones, afro-équatoriens et montubios illustre le résultat contemporain de décennies d'oppression linguistique et culturelle.

Actuellement, ces langues indigènes restent dévalorisées : même si elles sont devenues un patrimoine culturel et linguistique, ces langues sont minorisées et risquent de disparaître, car les nouvelles générations adoptent aujourd'hui la langue officielle, la langue d'une société globalisante qui les éloigne de leurs organisations communautaires. Même si la région amazonienne semble avoir conservé une grande partie de son patrimoine sociolinguistique, beaucoup de langues-cultures sont en danger d'extinction : le Sapara, pour ne citer qu'un seul exemple, ne possède pas plus de dix

---

<sup>2</sup> Disponible en ligne et consulté le 27 juillet 2022 : [https://www.oas.org/juridico/pdfs/mesicic4\\_ecu\\_const.pdf](https://www.oas.org/juridico/pdfs/mesicic4_ecu_const.pdf)

<sup>3</sup> « Le castellan est la langue officielle de l'Équateur ; le castellan, le kichwa et le shuar sont les langues officielles des relations interculturelles. Les autres langues ancestrales sont officielles pour les peuples indigènes dans leur territoire de vie et selon les termes que fixe la loi. »

locuteurs natifs. L'acteur social perd sa propre identité lorsqu'une langue étrangère lui est imposée comme langue officielle. Dans le cas de l'Équateur, cette réalité s'illustre par la domination de l'espagnol et s'exprime dans les discours produits par les producteurs de cacao : « La langue n'est pas seulement un moyen de communication, elle devient un moyen d'oppression » (Calvet, 1999 : 11).

Comme nous pouvons le constater à travers cette discussion, l'écologie du langage étudiée par de nombreux auteurs concerne la relation entre la langue et la réalité dans laquelle elle évolue incluant l'être humain, son environnement et les autres systèmes langagiers. La diversité linguistique fait partie intégrante de l'écologie des langues. En Équateur, quatorze langues-cultures se partagent un même environnement d'une grande biodiversité où chaque communauté socioculturelle décrit l'environnement à partir de sa propre richesse linguistique. Le fait qu'une langue dominante comme l'espagnol s'impose comme langue officielle de communication entraîne la disparition de langues indigènes et la perte de cultures autochtones.

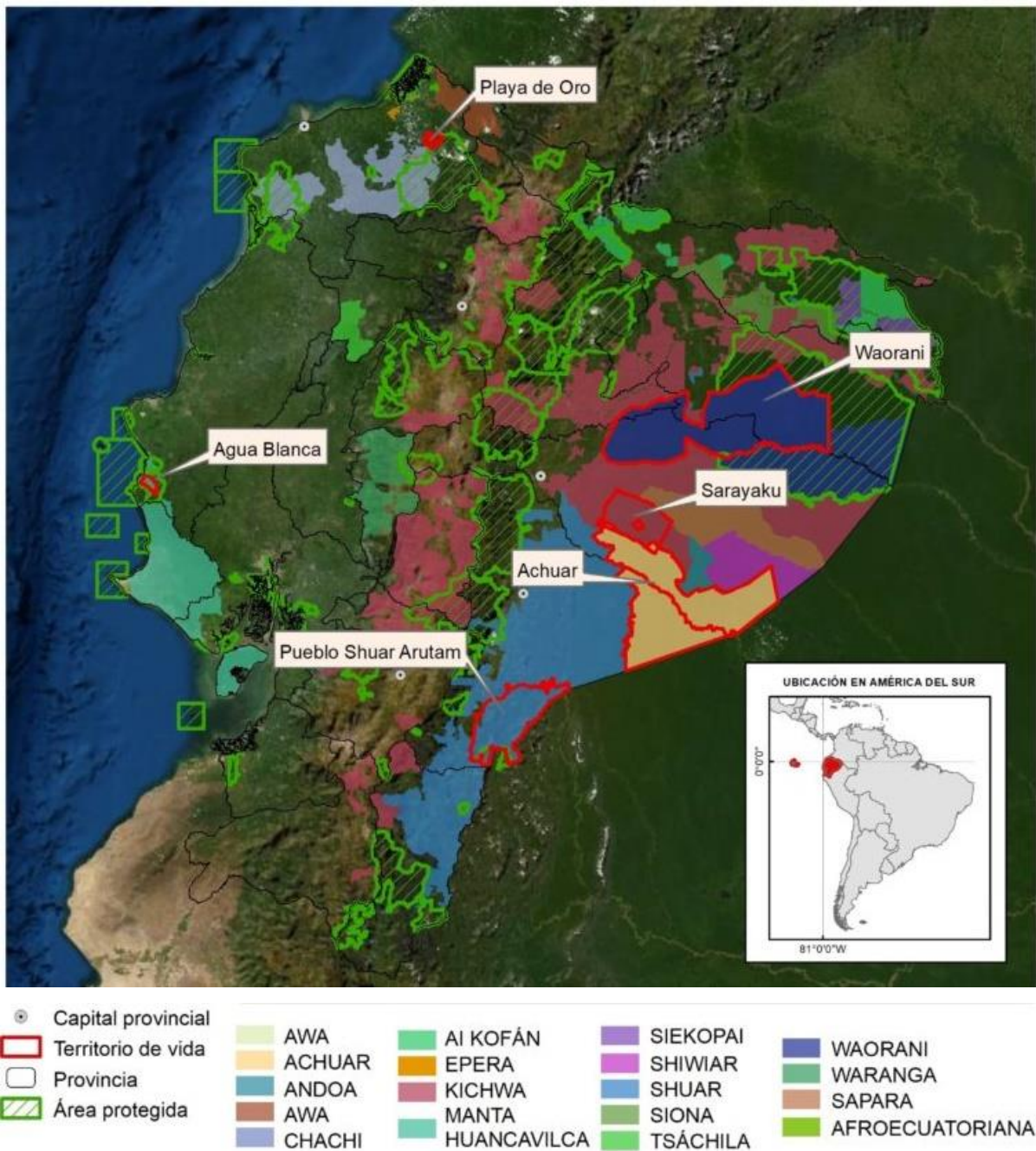


Figure 1 : Territoires des peuples et nationalités autochtones, afro-équatoriens et montubios

Source : Zamora G., Maldonado P. (2016) ; EcoCiencia, Atlas Amazonía Bajo Presión (2018) ; organisations des APAC en Équateur (2017-2020). Élaborée par : Fondation ALDEA, 2021.

Parmi les auteurs qui soutiennent la protection de la diversité linguistique et de la diversité culturelle, citons Mühlhäusler (1995), Harmond (1996) et Terralingua (2016),

ou encore le Programme des Nations Unies pour l'Environnement qui intègre la diversité culturelle humaine dans le concept de biodiversité considérant qu'elle peut être affectée par les mêmes facteurs que la biodiversité :

Le Sommet de Rio avait défini, par des textes juridiques de caractère contraignant, les liens entre les aspects socio-économiques, écologiques et culturels du développement. [...] La diversité culturelle reflétait la diversité naturelle et, chaque fois qu'une forêt était rasée et chaque fois qu'un langage parlé devenait une langue morte, une civilisation était amputée et un génocide était commis. (PNUE, 2002 : 3).

Il a été souligné que, dans la nature, la stabilité résulte de la diversité de ses composantes et de l'équilibre entre elles, de sorte que les écosystèmes fonctionnent en totale harmonie. La diversité des cultures, des valeurs et des civilisations et les interactions dynamiques entre celles-ci sont une source de stabilité dans le monde où nous vivons. [...] La mondialisation ne devrait pas avoir pour prix la perte de cultures et de traditions autochtones. Les valeurs spirituelles, la diversité culturelle, le respect des droits de l'homme et les connaissances autochtones peuvent contribuer utilement à la formulation d'une nouvelle éthique environnementale pour le XXI siècle. (PNUE, 2002 : 50).

Dans leurs rapports, les Nations Unies renforcent l'importance du lien entre diversité biologique et culturelle affirmant que la perte d'une de ces diversités conduit à la perte de la stabilité. Cette relation étroite entre diversité biologique et culturelle définit la richesse de l'écosystème linguistique considéré.

### 1.1.2. L'écologie du langage comme rapport entre langue et environnement naturel

L'écologie du langage apporte un cadre théorique général pertinent pour appréhender les changements du comportement communicatif selon l'écosystème dans lequel s'inscrit la situation de communication. Le fait même de mettre en relation l'écologie biologique et l'écologie linguistique ouvre un champ très vaste vers l'étude des écosystèmes linguistiques par analogie aux études d'écosystèmes naturels et de leur biodiversité dans un pays où la nature est au centre même de la cosmovision des communautés. En Équateur, l'écosystème visible - la réalité tangible - est divisé en cinq plans qui forme le *Sacha Runa Yachay* : *Awa pacha* qui correspond à l'espace-temps supérieur des astres, *Puyu pacha* correspond à l'espace-temps céleste des nuages, *Kay pacha* à l'espace-temps où nous vivons, *Uku pacha* à l'espace-temps de la terre d'en bas, et *Nina pacha* qui correspond à l'espace-temps du feu sous le *Uku pacha* (Andy Alvarado et coll. 2012 : 28, 117-129). Pour les communautés qui cultivent des écosystèmes aussi différents que ceux de la Cordillère des Andes, de l'Amazonie ou de la côte pacifique, dont la biodiversité a permis à l'Équateur d'obtenir son qualificatif de pays « mégadiver » (*pais megadiverso*), le respect de l'interrelation entre les cinq *pachas* s'exprime dans leur système de culture de produits alimentaires et médicinaux appelé *Chakra* dont la principale caractéristique est la polyculture et le respect de la biodiversité (Méric, 2018). Cependant, dès le début de la colonisation, cette cosmovision affronte des considérations allant des croyances religieuses jusqu'aux modèles d'économie de marché. Le modèle de la *Chakra* affronte un modèle de production intensif basé sur la monoculture. Le modèle communautaire centré sur la notion de *Sumak Kawsay*<sup>4</sup> affronte un modèle élitiste basé sur l'individu (Altmann, 2012 ; Jo, 2013 ; Caballero et Maldonado, 2016). Les langues-cultures autochtones affrontent l'extension d'une langue majoritaire véhiculant une économie de marché. Souffrant une colonisation toujours plus profonde, ces langues-cultures ancestrales sont

---

<sup>4</sup> Le *Sumak Kawsay* est un concept Kichwa qui se traduit en espagnol par « *Buen vivir* » et représente l'idée de « bien vivre » qui en français s'approche plus de la notion de « bien-être ».

aujourd'hui appelées minoritaires. Bastardas-Boada affirme que l'évolution conjointe entre langues minoritaires et langues majoritaires « est un processus intergénérationnel dirigé vers l'abandon du code minoritaire pour l'adoption de la langue majoritaire par la population qui utilisait auparavant le premier. L'analogie du gros poisson qui mange le petit, ou de l'extinction d'une espèce en raison de la destruction de son environnement naturel, vient facilement à l'esprit<sup>5</sup> » (2016 : 451-452). Il est aisé de comparer l'évolution de l'écosystème naturel équatorien : déforestation, contamination, perte de la biodiversité ; à l'évolution de l'écosystème du langage en Équateur : perte de locuteurs natifs de langues ancestrales, contamination terminologique de ces langues par les langues majoritaires, perte de la diversité des langues-cultures du pays comme le Sapara, déjà mentionné, qui ne compte pas plus d'une dizaine de locuteurs bien qu'en 2001 la culture Sapara ait été déclarée patrimoine immatériel de l'humanité. Cependant, l'auteur considère que « les langues ne sont ni des organismes ni des espèces biologiques, mais des comportements culturels humains contextuellement situés<sup>6</sup> » (2016 : 452), concept qui peut être mis en opposition avec l'approche de Mühlhäusler qui considère que le langage peut être étudié comme s'il s'agissait d'une relation entre des organismes et leur développement dans un environnement situé dans le temps et l'espace. En effet, Mühlhäusler considère l'apport de l'analyse écologique en termes de diversité linguistique et de colonisation, et il estime que l'approche écologique permet de comprendre et d'analyser la production du discours dans son cadre environnemental. En outre, il applique ses réflexions à tous les niveaux, de la lexicographie et de l'iconographie, et invite à une prise de conscience écologique en linguistique. Pour cet auteur, l'étude de l'écologie linguistique s'inspire de l'étude des êtres vivants ; c'est-à-dire de l'idée que le langage peut être étudié comme s'il s'agissait d'une relation entre

---

<sup>5</sup> Traduit par nous de « *su evolución presenta en muchos casos un proceso intergeneracional hacia el abandono del código minoritario y la adopción de la lengua mayoritaria por parte de la población que antes usaba el primero. Fácilmente viene a la mente la analogía del pez grande que se come al pequeño, o de la extinción de una especie a causa de la destrucción de su entorno natural* »

<sup>6</sup> Traduit par nous de « *las lenguas no son ni organismos ni especies biológicas sino comportamientos culturales humanos contextualmente situados* »

des organismes et leur développement dans un environnement situé dans le temps et l'espace. (Mühlhäusler, 1996)

Vivre cette évolution comme acteur social directement impliqué dans cette réalité ne peut qu'occasionner des changements et des variations qui transforment le discours, l'identité, la mémoire, et la perception de la réalité, car la cosmovision et les interprétations évoluent avec l'écosystème aussi bien naturel que socioculturel. La production nationale de cacao n'échappe pas à cette évolution. Sur la côte équatorienne, les producteurs privilégient un système de production intensif en monoculture sur de grandes extensions au détriment de la biodiversité naturelle. En Amazonie équatorienne, les producteurs cultivent de plus petites surfaces tout en observant un modèle de pluriculture comme ils le soulignent dans les entretiens (*Figure 2 : Zone de production du cacao*).



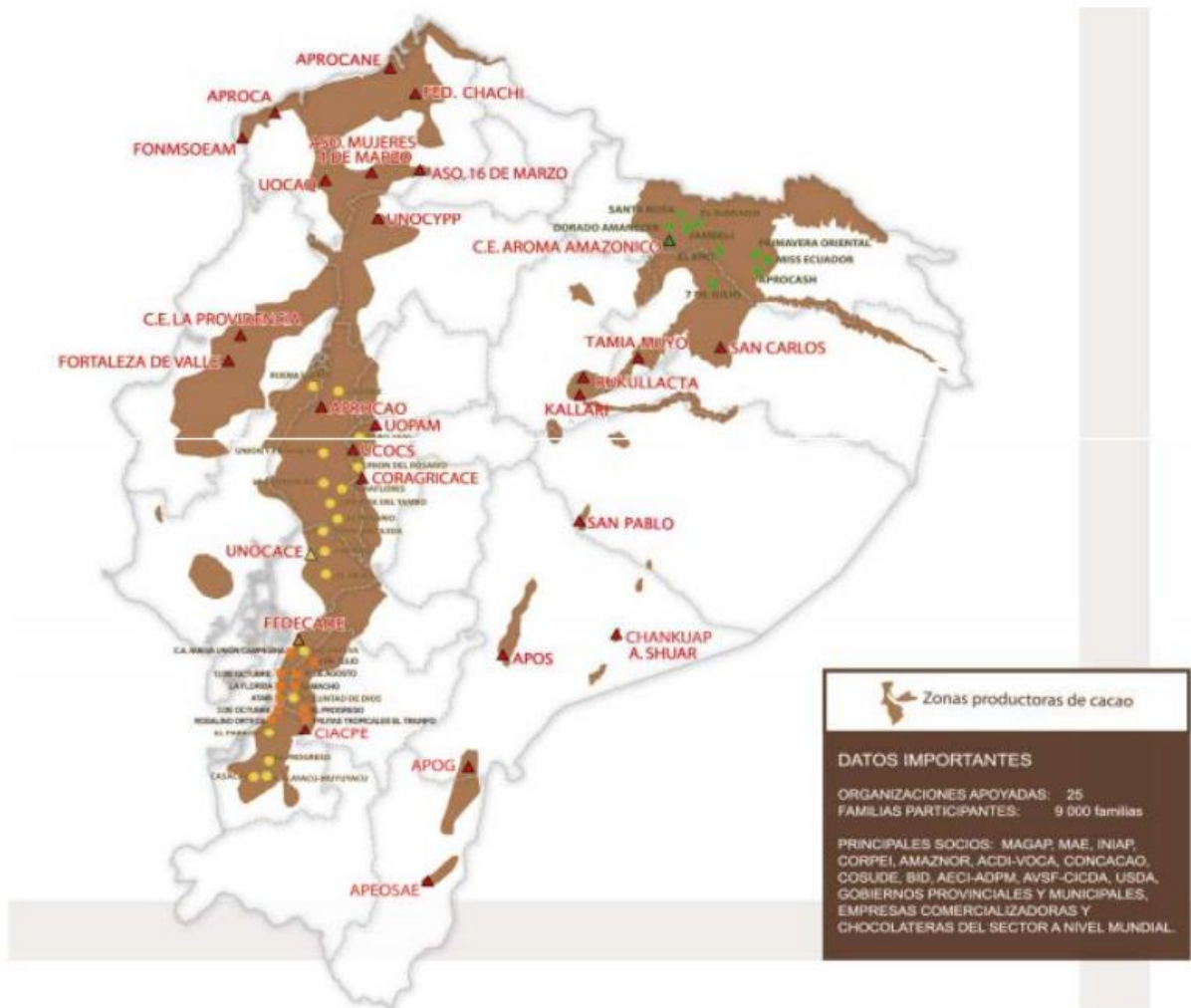


Figure 2 : Zone de production du cacao

Source : Enríquez Gustavo, CORPEI, Informe de Proyecto Mapa de Sabores, 2007.

Il est difficile de pouvoir apprécier pleinement l'information de la carte où sont situées les zones de productions du cacao sans avoir une idée de la géographie du pays. La Figure 3 : géographie de l'Équateur permet ainsi de situer la cordillère des Andes et de comprendre la division géographique du pays en trois régions continentales : la côte pacifique, la cordillère des Andes, et Amazonie ; et une région insulaire : les îles Galapagos.



Figure 3 : géographie de l'Équateur

Source : <https://www.mapas-del-mundo.net/america-del-sur/ecuador/grande-mapa-fisico-de-ecuador-con-carreteras-ciudades-y-aeropuertos>.

La côte, offrant un climat et une altitude propice à la culture du cacao près de grands ports d'exportation, est devenue l'environnement privilégié pour la production du cacao. Les Andes n'offrent pas de conditions propices à la culture du cacao, d'où la présence de cette bande blanche sur la carte de production du cacao (Figure 2 : Zone de production du cacao). L'Amazonie offre quant à elle d'excellentes conditions de

production pour le cacao, mais elle est éloignée des grands ports d'exportation, les coûts de production ne sont donc pas aussi compétitifs que sur la côte. Cependant, ils le deviennent pour des produits élaborés comme la barre de chocolat, d'où la naissance de structures de production du chocolat sous diverses formes comme nous le détaillerons dans la section 1.2.4. Chocolat actuellement produit en Équateur.

### 1.1.3. L'écologie du langage comme rapport entre langue et humains/société

Dans une approche socioculturelle de l'écologie du langage, le locuteur subit de nombreuses pressions de l'environnement dans lequel il s'exprime. Ces pressions ont un impact direct sur les stratégies de communication qu'il choisit, aussi bien dans la langue dominante que dans la langue dominée quand il appartient à une communauté pluriculturelle et plurilinguistique (Terborg, 2066 ; Terborg et García-Landa, 2013). Cette pression sociale s'exerce selon deux dimensions - le multilinguisme et le registre du statut social - lesquelles sont directement sous l'influence de la notion de pouvoir, de prestige et de l'intention de communication. Ces relations de pouvoir entre écosystèmes socioculturels définissent la vigueur, la transformation ou la disparition d'une langue-culture. Actuellement, il semble que les caractéristiques socio-économiques et les relations commerciales réunissent les principales variables qui déterminent l'usage et le devenir d'une langue-culture. En d'autres termes, l'écosystème dans lequel la langue s'inscrit subit les pressions politiques, sociales et économiques des écosystèmes en contact comme le souligne Bastardas-Boada :

Si un groupe d'humains entre fréquemment en contact étroit avec un autre, et que le premier est supérieur en nombre de locuteurs et en pouvoir politique ou

économique par rapport à l'autre, ses langues respectives peuvent alors se convertir en « majoritaire » et « minoritaire »<sup>7</sup>. (2016 : 452)

Les différences de pouvoir et de pression entre les groupes humains exercent une grande influence sur l'évolution de l'écosystème de la langue et donc son propre devenir. Cependant, l'avantage d'appréhender une langue par l'écosystème dans lequel elle vit, est d'offrir une perspective globale de l'« interrelation entre différents facteurs socioculturels et d'apprécier les dynamiques sociocognitives de la société<sup>8</sup> » (Bastardas-Boada 2016 : 452). Terborg, pour présenter son concept de l'écologie des pressions, étudie le déplacement des langues indigènes par l'espagnol, c'est-à-dire qu'il recherche les causes et les effets de ce phénomène. Le fait qu'un locuteur meurt sans laisser de locuteurs natifs peut être une des explications possibles, ou quand les langues entrent en contact et que l'une d'elles est déplacée jusqu'à son élimination complète. De ce rapport conflictuel apparaît le fait qu'une des langues en contact est déplacée par une autre. Les changements linguistiques qui se produisent dépendent des forces ou des pressions qui peuvent entrer en conflit les unes avec les autres. Dans l'écologie des langues équatoriennes, il existe différentes pressions pour que les langues indigènes soient déplacées par la langue officielle. La pression est liée aux relations de pouvoir, aux idéologies, aux valeurs, aux actions humaines et aux attitudes envers une certaine variante linguistique. Selon l'auteur, cette force de pression dépend de l'intérêt de l'individu, de sa volonté, et de sa résilience, lorsqu'il se sent sous la pression d'idéologies concurrentes (Terborg, 2006 : 2). Considérant la volonté politique, la Constitution définit directement le statut des langues en Équateur. Elle impose l'espagnol comme langue officielle. Elle positionne les langues ancestrales kichwa et shuar comme langues officielles de relation interculturelle, et relègue les autres langues ancestrales au rang de langues officielles pour les peuples ancestraux dans les zones où ils habitent. Dans la mosaïque des écosystèmes linguistiques de l'Équateur, la place de

---

<sup>7</sup> Traduit par nous de : « *Si un conjunto humano entra en contacto frecuente y estrecho con otro y el primero es muy superior en número de hablantes y en poder político y/o económico respecto al otro sus lenguas respectivas pueden convertirse entonces en « mayoritaria » una y « minoritaria » la otra »*

<sup>8</sup> Traduit par nous de : « *interrelación de los distintos factores y las dinámicas sociocognitivas de la sociedad »*

chaque langue-culture est déterminée sans équivoque ainsi que son prestige, en d'autres termes : son pouvoir. La conséquence directe d'une telle imposition est l'impossibilité pour une langue ancestrale de pouvoir changer son statut et donc de protéger celui de l'espagnol rendant utopique tout espoir de décolonisation sociolinguistique, car il faudrait imaginer une réforme constitutionnelle. La Constitution définit donc quelle est la langue majoritaire et quelles sont les langues minoritaires. La politique linguistique de l'état est claire : la conservation de la diversité linguistique ne peut pas se faire au détriment de la domination de la langue officielle. Cependant, la préoccupation pour la préservation de la diversité des systèmes linguistiques a créé la nécessité de comprendre en profondeur les mécanismes qui mènent à la substitution linguistique et, en fin de compte, à l'abandon total des langues minorisées (Junyent, 1991 ; Montoya-Abat, 1996). Cette réalité, aussi bien catalane qu'amérindienne, a suscité l'émergence d'un écologisme linguistique au service d'un activisme contestataire et d'une écolinguistique politique qui prétend pouvoir changer l'organisation socio-économique et culturelle des sociétés humaines. Accompagnant cet écologisme naît la proposition d'une égalité de droit des langues pour éviter les relations de subordination et de hiérarchisation (Junyent, 1991 ; Maffi, 2001 ; Bastardas-Boada, 2005 ; Skutnabb-Kangas et Phillipson, 2008 ; Bastardas-Boada, 2014) comme il peut être observé dans la Constitution équatorienne. Ce phénomène de conflit entre écosystèmes de langues différentes se superpose à un phénomène comparable qui, à moindre échelle, crée un conflit de pouvoir dans chaque écosystème de la langue où des variables socioculturelles comme la profession, le genre, le statut de l'acteur social ont un impact direct sur l'interprétation du discours produit. Les groupes ont du pouvoir s'ils sont capables de contrôler, dans leur propre intérêt, les actions et les esprits des autres groupes. Cette capacité présuppose un pouvoir de base consistant en un accès privilégié à des ressources sociales rares telles que la force, l'argent, le statut, la renommée, la connaissance, l'information, la culture ou même diverses formes de discours et de communications publiques selon les différentes intentions de communication. Toutes ces forces de pression, contradictoires ou non, font partie de l'écologie du langage, et se retrouvent profondément ancrées dans de nombreux discours. Les populations indigènes de



l'Équateur et leurs langues naturelles s'affrontent à une langue cultivée par le pouvoir politique et les exigences de rentabilité dans un écosystème de consommation. Actuellement, de nombreuses communautés essaient d'affirmer leur identité indigène et luttent pour leurs droits, pour la reconnaissance de leur patrimoine socioculturel incluant leur identité linguistique. Selon le statut du participant, cette réalité s'exprime dans les entretiens réalisés par les différences d'utilisation de la langue et le degré de variation linguistique que leurs productions langagières présentent par rapport l'espagnol considéré correct par l'Académie royale espagnole. Les producteurs de chocolat utilisent l'espagnol pour vanter les mérites de leurs produits et quand ils incluent quelques mots en langue ancestrale, leur intention est d'habiller leur produit d'une identité spécifique pour en faire la promotion. Il est même possible d'observer l'utilisation de l'anglais sur l'emballage de leurs produits vendus en Équateur démontrant ainsi que leurs chocolats respectent des critères internationaux (*Figure 4* : Étiquette de chocolat vendu en Équateur).



*Figure 4* : Étiquette de chocolat vendu en Équateur  
Source : Photo personnelle.

L'étiquette du chocolat *Minka* mélange du kichwa : *Shiwa Urku* (Coline de Palmier), de l'anglais : *Organic dark chocolate*, et de l'espagnol : *Hacienda* ou *Comunidad*, sans

oublier l'expression du poids en gramme et en once. Il est même possible de trouver certaines étiquettes avec des références françaises d'un chocolat fait et vendu en Équateur (*Figure 5 : Étiquette chocolat « au lait » vendu en Équateur*). L'intention de communication est dans ces cas clairement orientée vers l'utilisation d'écosystèmes de langue étrangers à ceux qui sont présents en Équateur pour que tout le poids culturel et référentiel de l'anglais et du français soit associé au produit vendu.



*Figure 5 : Étiquette chocolat « au lait » vendu en Équateur*  
*Source : Photo personnelle.*

Le discours du producteur de cacao dans la province de Pastaza semble plus étoffé dans leur langue naturelle que dans la langue cultivée même si dans leur production langagière en kichwa des termes espagnols sont utilisés principalement pour répondre à la technicité du processus de culture du cacao qui exige une formation qu'ils reçoivent exclusivement en espagnol. C'est ce rapport entre langue, humain et société qui caractérise l'écosystème linguistique dans lequel se produisent les discours compilés dans notre corpus d'étude.

#### 1.1.4 Discussion de la pertinence de ces approches pour le travail de thèse, de la construction et de l'analyse du corpus

En référence à la notion de régulation selon Calvet introduite précédemment, et par analogie à une approche darwinienne, la sélection sociale d'une innovation langagière, analogue au processus de sélection naturelle, et son acceptabilité par la communauté de langue, est mise en évidence par sa fréquence d'usage dans l'écologie du langage dans laquelle elle apparaît. Ces fréquences d'usage et autres variables statistiques ne peuvent être atteintes que par la compilation d'un ensemble de textes dans lesquels se matérialise le discours produit dans une situation de communication particulière située au sein d'une écologie de langue déterminée. Cet ensemble de textes constitue le corpus d'étude auquel les principes de la linguistique outillée s'appliquent de façon à obtenir les données statistiques objectives qui détermineront quelles réponses individuelles ont été socialement régulées et peuvent être considérées comme caractéristiques intrinsèques de l'expression de la sensorialité dans le discours spécialisé construit autour du chocolat en contexte équatorien. Il est important de noter que durant notre étude, dans la phase de collecte des données, il a été constaté que la grande majorité des membres des communautés indigènes de la province de Pastaza en Amazonie qui se consacrent à la culture du cacao parlent deux langues, la langue officielle (l'espagnol) et leur langue maternelle (le kichwa). Par conséquent, il existe un bilinguisme oral très marqué. Cependant, grâce à l'observation, nous avons pu nous rendre compte que leur langue maternelle est devenue un marqueur d'intimité dans leur environnement, dans leur groupe, dans leur entourage, dans une certaine mesure où elle est devenue un bouclier de protection contre la colonisation culturelle, l'inconnu, l'étrange. Elle est devenue un outil de revendication identitaire, même lorsque leur culture a été fortement influencée par la langue officielle.

Nous avons pu observer comment, lors du passage d'une langue à l'autre, le visage change, la fluidité s'améliore lors de l'expression et de la réponse aux questions. Nous



voyons comment le discours est conditionné, selon qu'il est exprimé dans la langue officielle ou dans la langue maternelle. Les acteurs sociaux considérés nous présentent des schémas d'organisation sémantique, phonétique ou syntaxique en accord avec leur réalité, leurs différentes identités, mais aussi avec leurs intentions de communication. Dans ces conditions, la langue officielle est utilisée comme outil de communication et d'intégration sociale au sein d'une écologie linguistique pluriculturelle afin d'avoir une meilleure position économique, éducative, etc. Même si les structures de leurs phrases ne sont pas correctes, s'ils ne coordonnent pas les mots, s'ils ne conjuguent pas les verbes, la langue officielle joue un rôle important pour le développement personnel et communautaire. Au même titre qu'une *lingua franca*, l'espagnol est devenu la langue de prestige, synonyme de réussite et d'intégration sociale. Elle est utilisée dans tous les domaines de la communication, reléguant les autres langues comme le shuar, l'achuar, le paicoca, l'awapit, le kichwa, entre autres, à l'expression d'une identité culturelle. La description de cette écologie linguistique est une étape nécessaire pour appréhender les interprétations qualitatives de résultats quantitatifs qu'une analyse de discours outillée peut apporter à partir d'un corpus de textes écolinguistiquement situés comme l'illustre cet exemple extrait du corpus :

*C-001<sup>9</sup> :<eh> bueno como como se puede decir bueno yo últimamente como osea no he hecho con mi con mi cacao pero así como como el ingeniero llegó él bueno ahí hemos cogido un tallercito por ahí y ahí va creo que hemos elaborado más que todo él.<sup>10</sup>*

Cette contribution est une réponse à la question : « Pouvez-vous définir les caractéristiques du chocolat fabriqué avec votre cacao ? ». Le producteur est gêné d'avouer qu'il n'a pas fabriqué de chocolat avec son cacao, qu'il ne se préoccupe en quelque sorte que de la production de la graine de cacao, sans réellement se soucier du

---

<sup>9</sup> Dans l'ensemble de ce travail, les exemples issus des contributions présentes dans le corpus compilé seront listés avec un numéro à trois chiffres précédés de la lettre C majuscule (initiale de contribution) et d'un tiret. La traduction de ces exemples sera proposée dans une note de bas de page et écrite en italique.

<sup>10</sup> <heu> bien comment comment dire bien moi ces derniers temps comment en fait je n'ai pas fait avec mon cacao avec mon cacao, mais quand l'ingénieur est venu lui bien on a participé à un petit atelier par là et c'est tout je crois qu'on en a fabriqué surtout lui (traduit par nous)

devenir de sa production malgré le *tallercito* (petit atelier) que l'ingénieur du ministère de l'Agriculture est venu animer en espagnol, et pendant lequel il pense avoir goûté le chocolat fabriqué à partir de son cacao.

Il est également nécessaire de mentionner que dans l'écologie linguistique équatorienne, le fait d'utiliser une langue officielle, et le fait que certains groupes maintiennent leur langue maternelle, créent un écosystème de langues en contact où il existe une influence réciproque des deux langues l'une sur l'autre qui, dans certains cas, peut aller jusqu'à la pratique d'un interlangage comme le kichwagnol comme l'illustre cet exemple extrait du corpus :

*C-002 : Bueno más que todo bueno ñuka mana kikinta tarpashkani, más que todo ñuka papito, ñuka finando viejo, pay ansa charishka chi tunuy, bueno chi tunuy sakikpi mejor ñuka ranti chi tunuy katishkani, anka chari como cinco, seis años chi tunuy ansa mantenimientota rasha katimushkani pero shinakllayta.<sup>11</sup>*

Lors de la phase de collecte des données des groupes d'étude, nous avons pu détecter une grande différence dans le discours utilisé par les producteurs impliqués dans la culture du cacao, par opposition à ceux impliqués dans la production, la commercialisation et la consommation du chocolat. L'exemple suivant met en parallèle le discours produit par un producteur de cacao et de chocolat quand ils répondent à deux questions de l'entrevue et s'expriment sur le produit fini : le chocolat.

<b>Producteur de chocolat</b>	<b>Producteur de cacao</b>
<b>entrevistador:</b> perfecto y ¿según usted cuáles son las características que definen a estos productos?	<b>entrevistador:</b> ya y puede definir las características del chocolate elaborado con su cacao
<b>entrevistado:</b> haber nosotros siempre que lo explicamos nosotros tenemos varias cualidades nosotros somos un producto orgánico somos un producto vegano es divertido no tenemos nada de leche ni trazas de leche ni nada es un producto vegano	<b>entrevistado:</b> bueno eso si no no no he no he hecho la prueba con mi propia mano no no no hecho no hecho
	<b>entrevistador:</b> no ha hecho la prueba

<sup>11</sup> Bien, en fait, moi je n'ai pas planté, en fait, mon père et mon grand-père décédé, ils m'ont laissé ce qu'ils avaient, donc j'ai continué, ça fait 5 ou 6 ans que je continue avec un entretien correct, je vous remercie.

toda la untuosidad que tienes del chocolate lo logra por la calidad del cacao fino de aroma  
somos producto kosher parve ok  
que elegimos con la certificación de calidad  
porque hay mucha gente que también lo consigue  
de hecho el parve te permite saber que tiene cero por ciento de de de lácteos  
porque para los judíos hay una época en en el año donde después de la pascua  
no puede tener nada que sea con ellos  
<eh> con el tema de de lácteos  
de ahí tienes un tema de  
tenemos de expresión orgánica americana y europea  
que realmente la fuerte fuerte de sacar la que quiere tea shoppes es la orgánica europea  
eso es un poco es eso  
pero aparte de eso ¿por qué?  
haber primero  
primero porque no solo haces chocolate  
tienes una historia para todos  
tienes historia de ayuda a la agricultura  
tienes una historia de ayuda mira  
puede ser desde los productores de cacao desde las productoras de uvilla  
nosotros les dimos a unas productoras de uvilla en la zona de Otavalo <eh> un deshidratador  
porque toda aquella fruta que no podían  
que no podías hacerla por la noche  
osea que no se podía venderla todo el día  
lo que se les iba perder  
nosotros les dimos un deshidratador  
para que deshidraten y enseguida nosotros se las compramos  
comenzamos con dos mujeres y ahora son cien familias  
**entrevistador:** interesante  
**entrevistado:** es un tema interesante  
es mucho más allá de vender chocolate

**entrevistador:** perfecto y ¿cuál cuál es la sensación que usted tiene cuando usted prueba este producto ?  
**entrevistado:** haber esto es un tema interesante  
todo depende de cual de ellos pruebas  
a mí el que más me gusta es Esmeraldas sesenta  
no tiene un sabor  
ese sabor chocolate sesenta por ciento  
no tiene ese amargor que uno dice chocolate negro  
debe ser amargo  
no es amargo

pero ha probado algún chocolate  
**entrevistado:** bueno si pero lo de afuera  
lo que hacen afuera eso si  
**entrevistador:** lo que hacen afuera  
y ¿qué sabor tiene este chocolate  
**entrevistado:** bueno como como estaba mencionó la la doña a veces como ponen con maní  
así con con otros sabores  
bueno ya distinto a ese sabor  
**entrevistador:** distinto sabor  
**entrevistado:** <aa>  
**entrevistador:** y cuándo probó  
cuándo prueba chocolate a qué le hace pensar?  
**entrevistado:** bueno yo no he yo no he  
bueno digamos alimentado para para ponernos  
se pensar en muchas cosas si no que  
ya en endulzar la boca nada más <jajaja>  
**entrevistador:** endulzar la boca nada más ya  
**entrevistado:** <aaa>

**entrevistador:** y y la sensación que le da cuando prueba <e>  
**entrevistado:** bueno eso si es bueno y  
como tiene ese saborcito y  
**entrevistador:** ya  
y las personas que que prueban este chocolate <ee>  
es decir un chocolate una clase de chocolate que que dice?  
¿qué ha escuchado decir usted?

es es untuoso es es muy rico  
de ahí tenemos el tema de  
tenemos sesenta setenta  
ochenta y cinco más azúcar de coco  
<e> cien y ciento uno  
si sabes como hacer ciento uno si  
son cien por ciento cacao más uno por ciento de nibs  
es un tema interesante  
es es untuosidad ese sabor  
ese sabor característico te recuerda al pasado a tu  
abuela  
a a compartir y además a que no ya te digo  
no tienes un tipo de barra  
tienes un  
es multifacético  
tienes además de todos los sabores  
que te puedes imaginar en todo latino américa

**entrevistado:** bueno una vez llegamos a un  
una fábrica de chocolates  
pero el el dueño mismo nos dijo que ese  
chocolate es así  
bueno como como tal estaba diciendo la  
compañera  
que hay que hay que trabajar  
con con bastante delicadez técnica  
**entrevistador:** ya  
**entrevistado:** <aa> nos dijo eso y  
de ahí nos hemos visto mismo  
no hemos llegado mismo a a la fábrica  
solamente lo que ellos han elaborado  
le han puesto entonces  
eso si compramos en polvo

Il paraît évident que leur terminologie ne présente pas les mêmes caractéristiques, que leurs intérêts ne sont pas les mêmes : celui qui cultive le cacao communique dans son discours qu'il aspire à une vie meilleure, à une bonne éducation pour ses enfants. Tandis que celui qui fabrique le chocolat communique avec un discours plus orienté sur le marketing, il cherche la reconnaissance de son produit et vante ses qualités. De la même manière, l'intention de communication du discours publicitaire est limpide : convaincre et vendre, contrairement à la production langagière de celui qui consomme qui exprime une expérience, le ressenti, la satisfaction de l'expérience sensorielle que procure le chocolat en bouche. Il apparaît évident que l'écologie linguistique propre à chaque acteur social, non seulement façonne son discours, mais oriente aussi son intention de communication. L'écologie de la langue offre donc un cadre théorique général pertinent qui permet au chercheur d'interpréter comment l'acteur social choisit ses stratégies de communication en fonction des caractéristiques de son environnement social, des variables socioculturelles comme le niveau d'éducation, la profession, et le lieu de production langagière.

## 1.2 L'approche ethnographique de la spécialité

La richesse de l'écologie du langage dans laquelle le discours spécialisé étudié se situe oblige à une approche analytique novatrice qui prend en compte dans ses fondamentaux théoriques l'interaction aussi bien culturelle que discursive mise en scène dans chaque situation de communication, ici spécialisée et professionnelle. Dans le contexte équatorien du domaine spécialisé où la fève de cacao locale se transforme en chocolat primé sur le marché international se côtoient autant d'interactions professionnelles qu'il existe d'experts, chacun intervenant à partir de ses différents écosystèmes linguistiques. Les interactions discursives entre professionnels évoluent dans une mosaïque socioculturelle aux identités si marquées que seule une approche ethnographique de la langue de spécialité peut permettre l'étude des discours comme composantes intrinsèques des compétences professionnelles de chaque niveau d'expertise (Wozniak, 2019 : 47). En effet, selon Wozniak, l'approche empirique et ethnographique de la spécialité permet d'aborder la composante discursive de l'expert à partir des interactions produites dans des situations professionnelles situées (2019 : 53-54), pour notre étude, dans la catégorie des activités socio-économiques, et d'envisager que c'est bien le domaine de spécialité qui fait le discours spécialisé, et non la terminologie spécialisée qui construit la spécialité de la discipline (2019 : 17). En d'autres termes, le degré de spécialité de l'écosystème linguistique détermine le degré de spécialité du discours. Dans sa monographie, Wozniak propose un modèle pas-à-pas d'analyse de caractérisation des domaines spécialisés (2019 : 87-88) et de leurs discours : une fois la problématique de recherche définie, elle préconise la découverte de la spécialité en contexte, dans son cas anglophone et francophone, afin de modéliser le discours de spécialité dans le respect de son écologie. Pour notre étude, l'écologie du discours du sensoriel dans le domaine de la production du chocolat s'articule autour de l'expérience, sociale autant qu'individuelle, et des émotions (Bach, 2021b : 107). Ce discours ne peut donc être mis au jour et reconstruit qu'au travers de son domaine d'activité située dans un écosystème dont la principale intention de communication est de partager aussi bien

une expérience qu'une émotion. « L'environnement influence la cognition et la langue, et réciproquement ; puisque la perception évolue selon l'environnement, le résultat final de cette perception est aussi influencé » (Bach, 2021b : 109). La perception cognitive d'un carré de chocolat dégusté dans une plantation de cacao en Équateur ne sera pas comparable à la perception cognitive d'un carré de chocolat identique dégusté par le même individu lors d'une opération de promotion dans un supermarché français. L'environnement étant différent, le résultat de cette perception et son expression discursive le seront aussi.

Adaptable au terrain équatorien, la complémentarité d'une approche ethnographique et d'une linguistique cognitive et outillée offre une caractérisation pertinente des différents discours produits. L'observation sur le terrain guide le chercheur vers une compréhension juste et pertinente de l'écologie de la langue. La constitution d'un corpus et l'analyse, elle-même dirigée et orientée par le corpus, conduisent à l'interprétation du processus de construction de la perception sensorielle à partir de l'expérience cognitive présente et passée (Bach, 2021b : 107-108).

En résumé, l'importance de l'approche ethnographique réside dans l'analyse systématique de la culture des groupes humains afin de connaître la réalité des peuples autochtones. Pour recueillir les informations, il est souvent nécessaire de gagner la confiance de la communauté, en plus de comprendre la culture et le monde qui l'entourent. Il est donc nécessaire d'analyser d'un point de vue interne et externe en utilisant des techniques telles que l'observation, les entretiens et les enquêtes personnelles afin de compiler l'information recueillie au sein d'un corpus représentatif.

Grâce à cette méthodologie, nous avons pu collecter les informations pertinentes pour nous approcher de la réalité sociale des communautés de la province de Pastaza, ainsi que des experts producteurs de chocolat. De cette façon, nous proposons une interprétation et une lecture des interactions professionnelles de chaque expert dans lesquelles ils construisent un discours pour exprimer leurs différentes perceptions sensorielles.

### 1.2.1. Représentation sociale du cacao et du chocolat en Équateur

Reconnu pour l'excellence des saveurs et des arômes de son cacao fin plus connu sous le nom de *Arriba*, l'Équateur est un pays dont l'histoire a été marquée par la production de cette fève au point d'en devenir un symbole identitaire. La *pepa de oro* entre dans la production des meilleurs chocolats fabriqués dans le monde entier participant ainsi à la fierté nationale d'une activité économique, sociale et culturelle incontournable :

*C-003 : entrevistador: ¿Cuáles son las palabras más utilizadas para hablar de este chocolate?*

*entrevistado: Ya aquí más que todo nos tratamos o nos decimos que cacao es pepa de oro. A nivel nacional ya dicen que es pepa de oro. Entonces a nosotros ennoblece nos orgullece con eso que tenemos<sup>12</sup>*

Le témoignage de cet agriculteur démontre que l'activité cacaotière en Équateur est présente dès les premières étapes de la fabrication du chocolat, c'est-à-dire dès la production de la fève, qui même si elle reste dans la plupart des cas une activité familiale, participe à l'économie de tout un pays. Les changements qui se sont produits au cours de l'histoire ont permis le développement de pratiques et de techniques de production et de transformation en chocolat. Cependant, il existe une forte segmentation dans l'organisation du processus de production du chocolat. Les agriculteurs vendent la fève sèche aux professionnels de la transformation qui produisent le chocolat d'exportation aux qualités internationalement primées. Il est rare de rencontrer des personnes qui interviennent dans l'intégralité du processus, de la plante à la barre de chocolat comme c'est régulièrement le cas en France pour la production du vin. Le système général de la filière s'approche plus de l'organisation de coopérative

---

<sup>12</sup> **intervieweur** : Quels sont les mots les plus utilisés pour parler de ce chocolat ? **interviewé** : Bien, ici, le plus souvent, nous utilisons ou nous disons que le cacao est une pépite d'or. Mais maintenant, c'est au niveau national, qu'ils disent que c'est une pépite d'or. Ce qui nous donne nos lettres de noblesse, nous rend fiers de ce que nous avons. (traduit par nous)

d'exportation ou d'entreprises qui fabriquent le chocolat. Ces structures achètent la fève de cacao aux petits producteurs pour être soit exportées, soit transformées sur place. Ces organisations dominent la filière et ne laissent que peu de marge d'action aux cultivateurs de cacao qui, pour la plupart, ne connaissent pas la saveur du produit fini. Pour avoir une référence européenne, il est possible de comparer l'organisation de la filière du cacao en Équateur à celle de la filière vitivinicole de la Rioja (Brémond, 2011) où 96 % des barriques sont dominées par le secteur élaborateur, les *bodegas criadoras*, qui se chargent d'affiner le vin, de le commercialiser et d'en faire la promotion exerçant une domination sur les petits viticulteurs.

Cependant, il existe des entrepreneurs qui ont choisi de s'impliquer dans l'ensemble du processus de fabrication du chocolat en développant des stratégies de vente liées au tourisme comme il est possible d'en rencontrer sur les routes du cacao de la côte et de l'Amazonie équatorienne :

*C-004 : entrevistado: Este chocolate nos preparamos más que todo a esencia de vainilla, con la esencia de ajengibre, de anís, canela, limón. Más que todo eso le gusta a los visitantes, más que todo a los turistas extranjeros.<sup>13</sup>*

Cette forme naissante d'agrotourisme permet aux visiteurs de fabriquer leur propre chocolat de manière artisanale, d'exposer leur savoir-faire et de créer un environnement professionnel où le discours spécialisé, et notamment celui relatif à la sensorialité dans le cas du chocolat, trouve une place privilégiée de production entre professionnels et *aficionados*. Cependant, dans l'écosystème linguistique équatorien, il reste encore à définir les compétences langagières liées à la spécialité ainsi que le cadre formatif pour un discours spécialisé afin d'améliorer les compétences professionnelles discursives et transversales liées à la communication située dans un domaine spécialisé (Laygues, 2018). Même si ces initiatives sont de plus en plus nombreuses, elles ne

---

<sup>13</sup> **interviewé** : Ce chocolat, nous le préparons le plus souvent avec de l'essence de vanille, avec de l'essence de gingembre, d'anis, de cannelle, de citron. C'est ce qui plaît le plus aux visiteurs, ce qui plaît le plus aux touristes étrangers. (traduit par nous)



représentent pas le quotidien de la majorité des producteurs de cacao qui, en Amazonie, vivent au sein de communautés difficiles d'accès où leur modèle d'agriculture suit le modèle d'organisation traditionnel de la *Chakra* (Méric et Carrera, 2018). Ce sont des parcelles de terrain où est pratiquée la polyculture, où le cacao se retrouve au milieu d'autres produits comme la yuca (parent du manioc africain), les bananes, les citrons et autres produits amazoniens qui ne se récoltent pas à la même époque. Cette organisation communautaire est propre à la culture des différentes ethnies qui le plus souvent ont l'espagnol comme deuxième langue, c'est le cas des kichwa de l'Amazonie. Ils cultivent des parcelles de l'ordre de l'hectare, vendent leur production sans aucun processus de transformation ; ils le cultivent, le récoltent, le séchent et le vendent. Certains n'ont jamais eu l'occasion de goûter le chocolat fabriqué à partir de leur produit qu'ils appellent *pepa de oro* sans vraiment en connaître la valeur :

*C-005 : entrevistador: ... puede definir puede decirme las características que tiene este... el chocolate que usted elabora con su cacao?*

*entrevistado: es que no no hecho... no he probado con eso... nunca he estado con con un poco de...<sup>14</sup>*

Face à cette réalité, il existe des initiatives du ministère de l'Agriculture équatorien pour former les agriculteurs à l'amélioration des techniques de production de la fève de cacao, mais aussi du chocolat artisanal. Grâce à ces ateliers de formation, nous avons pu accompagner les ingénieurs responsables de ces cours, ce qui a facilité l'accès, le contact et la réalisation des entrevues au sein de communautés généralement hermétiques à la présence d'étrangers même équatoriens. Cependant, une fois acceptés, les gens sont très généreux, ils nous ont accueillis de façon amicale et nous ont offert de partager leur quotidien : préparation d'un plat typique de la région appelé *Maito de carachama*, boissons chaudes à partir de l'écorce des arbres (*chuchuwuaso*), préparation de la *chicha* (boisson fermentée à base de manioc) et participation aux

---

<sup>14</sup> **intervieweur** : ...vous pouvez définir... vous pouvez me dire les caractéristiques que possède le chocolat que vous faites avec votre cacao ? **interviewé** : c'est que, non, je n'ai pas fait, je n'ai pas goûté comme ça, je n'ai jamais eu un peu de... (traduit par nous)

diverses activités qu'ils réalisent dans la Chakra, le tout dans un esprit tout à fait communautaire où être accepté implique d'accepter tout ce qu'il vous offre.

### 1.2.2. Description du cacaoyer et de son fruit

Selon le dictionnaire de l'Académie royale espagnole, le cacaoyer est un « arbre américain, de la famille des *sterculiacées*, au tronc lisse, haut de cinq à huit mètres, aux feuilles alternées, brillantes, lisses, dures et ovales, aux petites fleurs jaune et vermillon, dont le fruit est utilisé comme ingrédient principal du chocolat »<sup>15</sup>.

En effet, le cacaoyer ou cacaotier est une plante tropicale, qui a entre quatre et huit mètres de haut, de la famille des *Sterculiacées*, dont le nom scientifique est *Théobroma cacao*, qui signifie en grec : nourriture des dieux. La tige du cacaoyer est lisse avec une écorce sombre, gris-brun. Ses feuilles sont ovales et légèrement asymétriques, ses couleurs sont variables du violet au vert pâle. Ses fleurs se présentent en grappes le long du tronc et des branches, la fleur est vermillon, violet et blanc, en forme d'étoile. Le cacaoyer produit un fruit ligneux en forme de cabosse allongée de couleur rouge, blanche ou verte qui fonce en mûrissant, le fruit ou la cabosse mesure entre 10 et 32 cm de long et entre 7 et 10 cm de large et pèse entre 200 gr et 1 kg. À l'intérieur, il contient entre 20 et 60 graines, disposées en 5 rangées entourées d'une pulpe gélatineuse et sucrée, appelée le mucilage du cacao. Il pousse dans les climats équatoriaux où les précipitations sont abondantes tout au long de l'année et où la température est stable entre 25 et 28 degrés Celsius<sup>16</sup> (Alcívar Rodríguez, Pino Peralta et Aguilar Azuero, 2019).

La cabosse de cacao est composée de trois parties : l'écorce, les fèves et le mucilage ; son écorce rugueuse, d'une épaisseur d'environ 3 cm, contient le mucilage et les fèves

---

<sup>15</sup> Texte de l'académie royale espagnole traduit par nous disponible : <https://dle.rae.es/cacao?m=form>

<sup>16</sup> Traduit par nous du livre : « *El gran cacao de los pequeños productores* » (2019)

de cacao, et représente environ 64 % du poids total de la cabosse (Cadena Fiallos, 2015). Elle contient un axe central appelé placenta auquel sont attachées les fèves de cacao et une partie de la pulpe dont l'une des principales fonctions est de fournir des nutriments aux fèves de cacao d'où son nom de placenta. Elle représente également environ 5 % du poids total de la cabosse. Les graines ou les fèves de cacao sont enveloppées d'un liquide appelé mucilage, qui est composé de cellules de sève riches en sucre à 10-13 %, en acide citrique à 1-2 %, en pentoses à 2-3 % et en pectine à 8-10 % (Yépez, 2015). La racine principale atteint jusqu'à 3 m de long (Luzuriaga, 2012).

De nombreuses études vantent les bénéfices du cacao, et par conséquent du chocolat. Une simple recherche de « health benefit of cocoa » sur Internet démontre qu'il est considéré comme un puissant antioxydant grâce aux nombreux polyphénols qu'il contient (Jean-Marie et coll., 2021). Les antioxydants diminuent le risque de maladies cardiaques et agissent sur les dégâts produits par le cholestérol tout comme les huiles essentielles monoinsaturées oléiques qui se retrouvent dans l'huile d'olive. Le chocolat possède des nutriments comme les hydrates de carbone, l'acide folique, ou encore des minéraux comme le cuivre, le fer, le zinc et le magnésium qui participent au renforcement du système immunitaire.

## **Types de cacao**

L'ancienne classification du cacao se concentrait sur trois types : *criollo*, *forastero* et *trinitario* (un mélange de *criollo* et de *forastero*).

1. **Cacao *Criollo*** : les amandes sont d'un blanc ivoire, ses épis sont allongés, verts et rougeâtres à l'état non mûr, lorsqu'ils sont mûrs, leur couleur devient jaune et orange-rougeâtre. Ce chocolat est apprécié pour son goût de noisette et de fruit, il est considéré comme un cacao fin.
2. **Cacao *Forastero*** : les épis sont verts avant de mûrir, puis ils deviennent jaunes, les amandes sont aplaties et petites, ce type de cacao a une saveur basique.

3. **Cacao de *Trinitario*** : ce cacao est le résultat d'un croisement entre le *Criollo* de Trinidad et le *Forastero*. Il a une saveur de cacao moyenne élevée, généralement accompagnée d'une saveur de fruit et d'amande.<sup>17</sup>

## Types de cacao en Équateur

Selon le ministère de l'Agriculture, de l'Élevage et de la Pêche équatorien (MAGAP) et les recherches effectuées en Équateur, nous avons deux types de cacao :

1. **Cacao *Nacional o Arriba*** est considéré comme le produit emblématique de l'Équateur, reconnu dans le monde entier pour ses caractéristiques uniques en termes de finesse d'arôme et de saveur qui sont dues aux conditions climatiques et géographiques de l'Équateur. Les fèves *Arriba* sont sélectionnées pour élaborer des chocolats fins de grande valeur destinés à des segments du marché plus exigeants.

L'Équateur produit 63 % de la production mondiale du cacao aromatique dit *fino*, à partir de son produit endémique appelé *Nacional* ou *Arriba*, qui est recherché et préféré par les meilleurs chocolatiers du marché international.

Les caractéristiques organoleptiques ont permis une différenciation au niveau international, reconnue par l'industrie de la confiserie du secteur producteur de cacao. Caractérisé par une « fermentation très courte et donnant un chocolat lisse avec une bonne saveur et un bon arôme, il est identifié internationalement avec la classification de cacao à arôme fin » (Quingaísa et Riveros, 2007 : 4). C'est avec ce cacao que l'on obtient un des meilleurs chocolats du monde, pour sa saveur et son arôme floral, associé à des profils fruités et à d'autres saveurs.

2. **Cacao CCN-51** est une variété de cacao créée par l'homme, identifiée par un code alphanumérique, un cacao qui couvre actuellement une partie des

---

<sup>17</sup> Traduit par nous du livre : « *Manual de cultivo de cacao para la amazonía ecuatoriana* » (2009), <http://canacacao.org/wp-content/uploads/Manual-cultivo-Cacao-Ecuador.pdf>

plantations de l'Amazonie. Il se distingue par la couleur de ses cabosses qui est plus rougeâtre dans le stade de sa maturité, elles contiennent de plus grandes quantités de matières grasses et de mucilage. Ce cacao est généralement préféré pour la fabrication de chocolat en gros et autres produits finis dérivés comme le beurre de cacao.

Les méthodes d'amélioration du processus de production du cacao permettent d'augmenter les caractéristiques sanitaires et la productivité de la plante, ce qui assure sa qualité et la possibilité de conserver les propriétés du cacao *Arriba* reconnues par les experts internationaux. La sélection de clones obtenus à partir de souche mère *Trinitario* et *Nacional* présente des caractéristiques hautement compétitives de lignes hybrides interclonales de cacao, résultat de méthodes de sélection qui ont généré des clones très résistants et productifs, tout en conservant les caractéristiques aromatiques (Vera et Goya, 2015). De nombreuses stratégies issues de différentes recherches, publiées aussi bien par des universitaires que par des experts professionnels, permettent d'apporter des améliorations au processus de production du cacao et plus récemment au processus de fabrication du chocolat.

### 1.2.3. Histoire du cacao en Équateur

L'Amérique latine est connue comme le berceau du cacao. Des études menées en 2007 affirment que l'une des variétés de *Theobroma cacao*, est originaire de l'Amazonie équatorienne, dans l'actuelle province de Zamora Chinchipe au sud-est de l'Équateur, les vestiges trouvés dans cette région ont permis d'affirmer que le cacao était connu et utilisé dans les cultures préhispaniques.

On dit traditionnellement que cette plante a été utilisée pour la première fois en Méso-Amérique. La preuve la plus ancienne du Mexique provient du site de Mokaya, situé dans le col d'Almada, sur la côte Pacifique sud du Mexique,

les dates associées vont de 1900 à 1500 avant J.-C. D'autres dates précoces d'utilisation du cacao proviennent de plusieurs sites olmèques. À el Manatí, les dates associées vont de 1650 à 1500 avant J.-C. À San Lorenzo, les dates se situent entre 1800 et 1000 avant J.-C. (Powis et coll., 2011). À Palanda, il a été prouvé que son utilisation remonte à plus de 1500 ans. (Valdez, 2013 : 53-55)

Cependant, Valdez, dans sa publication *Primeras Sociedades de la Alta Amazonía*, décrit les résultats de ses fouilles archéologiques qui ont permis de faire remonter l'origine du cacao et de sa préparation à plus de 5000 ans dans les terres de l'Amazonie équatorienne. Cette variété de cacao est actuellement connue sous le nom de cacao *fino de aroma* (2013 : 53-55).

Le sujet de l'origine du cacao a été débattu pendant des décennies, de nombreuses enquêtes historiques selon Páez Paredes font référence à l'utilisation du cacao dans l'antiquité : les peuples olmèques, mayas et aztèques préparaient une boisson amère appelée *tchocolatl*, qui était considérée comme un aliment sacré pour l'empereur, la noblesse et les guerriers. Une des légendes mexicaines raconte une anecdote sur la rencontre de l'empereur aztèque Montezuma avec le conquérant, Hernán Cortés, lorsqu'on lui a offert le *tchocolatl*, une boisson énergisante destinée aux cérémonies spéciales. Cortés a été très surpris, à cause du goût fort et de l'épaisse couche de mousse. Il a été encore plus surpris lorsque l'empereur et ses généraux ont bu avec révérence. Il ajoute que les conquérants espagnols ont saccagé les caves royales à la recherche d'or et de pierres précieuses, mais qu'ils ont été surpris de trouver des millions de graines stockées en toute sécurité, car ces graines étaient un trésor précieux, le cacao étant utilisé comme moyen d'échange de biens et de services, « la fève de cacao était utilisée comme monnaie d'échange » (Páez Paredes, 2015 : 23), elle fut appelée « monnaie primitive ».

Au XVI<sup>e</sup> siècle, la boisson *tchocolatl* commence à être connue, appréciée et devient la boisson des aristocrates. Peu à peu, elle acquiert du prestige en tant que stimulant et se

répand dans toute l'Europe. Au milieu du XVIe siècle, le cacao est devenu une activité rentable, suscitant l'intérêt des hommes d'affaires de Guayaquil qui ont décidé de cultiver ce produit, malgré les interdictions établies par les décrets royaux. Les *guayaquileños* ont commercialisé clandestinement ce produit. En 1789, le roi Carlos IV a autorisé, par décret royal, la culture et l'exportation du cacao de la côte équatorienne. En 1789, les exportations de Guayaquil vers l'Europe et les États-Unis ont commencé, devenant ainsi le premier produit d'exportation équatorien.

Au cours de la seconde moitié du XIXe siècle, la production de cacao a connu plusieurs pics de développement. Cependant, c'est au cours de la décennie de 1890 que la production de cacao a connu sa plus importante progression, qui a marqué le début du boom cacaoyer (explosion du marché du cacao). Au cours de la période 1891-1920, la balance commerciale de l'Équateur a été excédentaire pendant vingt-neuf des trente années considérées. L'apparition de nouvelles fabriques de chocolat, notamment sur le marché américain, a augmenté la quantité de cacao importé vers ce pays et aussi d'Europe<sup>18</sup>.

En 1907, l'Équateur est entré dans une période de crise du cacao due à des maladies telles que la *monilla* et le balai de sorcière (Porrás et Sánchez, 1991) transmises par différents parasites de la plante. En 1914, avec la Première Guerre mondiale, la consommation de cacao diminue, les prix baissent et le secteur des producteurs et exportateurs du cacao, ainsi que l'ensemble du système économique équatorien, est lourdement affecté en faisant entrer l'Équateur dans une période d'instabilité.

Le « boom du cacao » a permis à des milliers d'Équatoriens de s'impliquer dans la production de la *pepa de oro*, le plus souvent comme ouvriers, travailleurs, journaliers laissant un héritage pragmatique des pratiques et savoirs liés à la culture, à la récolte et à la post-récolte du cacao. Ils ont aussi participé à ancrer le discours du domaine spécifique de la production du cacao dans le paysage écolinguistique de l'Équateur. Cet

---

<sup>18</sup> Pour de plus amples informations sur le sujet, veuillez consulter l'article du « *Foro Economía Ecuador* » (<http://foroeconomiaecuador.com/fee/los-booms-en-perspectiva-cacao-banano/>)

héritage s'est transmis de génération en génération dans de nombreuses villes de la côte équatorienne comme : Quevedo, Mocache, Palanque, Vinces, Pueblo Viejo, Bahía de Caráquez, Naranjal, Machala, Chone, Jambambalaya et Chone, Jambelí, Esmeraldas et Babahoyo, dont l'économie a fortement été dépendante du marché international de l'import-export de la fève de cacao<sup>19</sup> lui donnant ses lettres de noblesse et le surnom de *pepa de oro*.

La plus grande partie du cacao équatorien correspond à un mélange de *Nacional* et de *Trinitario* introduit après 1920 parce qu'il est considéré comme plus résistant aux maladies. Cependant, la saveur de la variété *Arriba* demeure, car l'Équateur possède les conditions agroclimatiques nécessaires pour le développement de cette production. Cela signifie que la qualité du cacao équatorien est due à la richesse du sol et du climat, ainsi qu'à sa situation géographique équatoriale. Le cacao équatorien est reconnu mondialement comme un leader dans la production de cacao *fino de aroma*, grâce à cette variété de cacao « *Arriba* ». L'aspect socio-économique de l'Équateur a longtemps dépendu de la production et du marché international du cacao, car il s'agit d'un produit symbolique du pays, faisant ainsi la différence avec les autres pays producteurs de cacao (Pazmiño Arízaga, 2015).

Selon la population des zones cacaotières équatoriennes, « le cacao est la tirelire du pauvre ». Cette représentation sociale semble, selon les observations faites sur le terrain, être bien adaptée à la réalité socio-économique des habitants des communautés de la province de Pastaza qui soulignent que c'est une source importante de revenus pour survivre et éduquer leurs enfants. Le cacao est un des produits privilégiés de la production agricole de Pastaza au même titre que la canne à sucre, la banane plantain ou la yuca.

Selon l'institut national de statistique de l'Équateur, le secteur cacaoyer contribue à hauteur de 5 % de la population économiquement active et occupe la troisième place

---

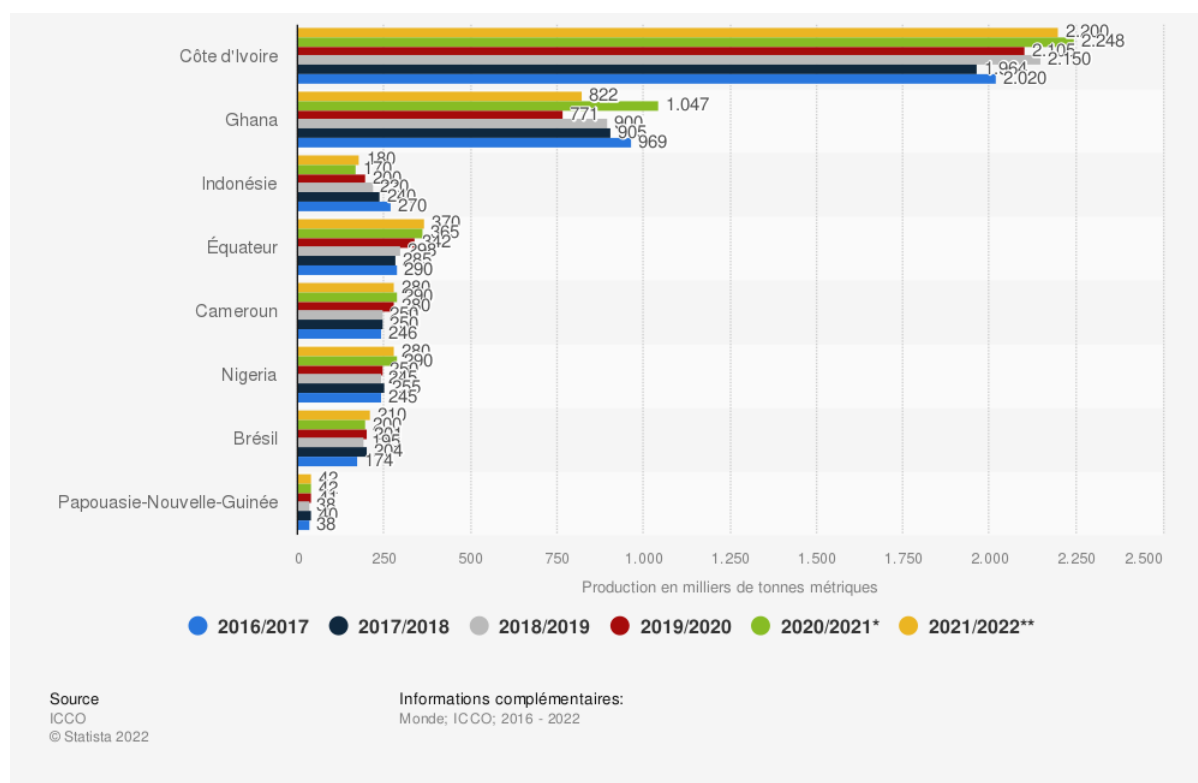
<sup>19</sup> Pour de plus amples informations sur le sujet, veuillez consulter l'article du « *Cacao Ecuatoriano Historia e identidad* ». <http://ahoranews.net/cacao-ecuatoriano-historia-e-identidad/>



sur le podium des producteurs avec 7 % de la production mondiale de cacao sachant qu’avec le pétrole et le café, le cacao est l’un des trois principaux produits négociés sur les bourses mondiales de matières premières.

#### 1.2.4. Chocolat actuellement produit en Équateur

Les 370 milliers de tonnes qui correspondent à la production de l’année 2021 et la projection de l’année 2022 positionnent l’Équateur comme troisième pays producteur de cacao sur le marché international (*Figure 6 : Production mondiale de fèves de cacao en volume de 2016/2017 à 2021/2022*).



*Figure 6 : Production mondiale de fèves de cacao en volume de 2016/2017 à 2021/2022*  
 Source : <https://fr.statista.com/statistiques/565101/production-mondiale-feves-cacao-volume-par-pays/>

Actuellement, l'Équateur conserve deux espèces de cacao produites grâce à des conditions agro-climatiques favorables qui permettent au cacao national ou Arriba et au CCN-51 (hybride) de présenter une excellente qualité.

Aujourd'hui, le cacao équatorien a retrouvé sa splendeur, l'Équateur étant reconnu comme le plus grand producteur et exportateur de cacao *fino de aroma* au monde. Nous constatons une forte transformation sociale et culturelle à travers le discours que les producteurs et vendeurs de chocolat alimentent pour positionner leur produit dans un marché international haut de gamme. Les aficionados - non-experts - suivent cette transformation en essayant de s'approprier le discours de l'expert dans leurs productions langagières. Cependant, le discours des producteurs de la fève de cacao semble exclu de cet écosystème élitiste. À la lecture des transcriptions des enregistrements réalisés, le lecteur découvre intuitivement des différences discursives entre les producteurs de cacao qui vendent leur produit, les fabricants et les vendeurs de chocolat, et les consommateurs du produit. Ces différences mettent en évidence la diversité des écosystèmes impliqués dans la chaîne de production du chocolat. Cette étude nous permettra d'analyser et de décrire ces spécificités et cette variété discursive intuitivement perçues.

L'Équateur, en tant que l'un des principaux producteurs de cacao *fino de aroma* au monde, a commencé à investir dans la transformation de ce cacao en chocolat aussi bien chez les grands que chez les petits producteurs pour éviter de simplement exporter la fève. De nombreux entrepreneurs industrialisent le cacao, en fabriquant des produits élaborés et semi-élaborés qui se sont positionnés au plus haut niveau dans la sphère internationale des meilleurs producteurs de chocolat au monde<sup>20</sup>. La qualité reconnue et appréciée du cacao équatorien a permis à l'Équateur de se voir décerner des prix internationaux tels que ceux que nous pouvons citer ci-dessous :

---

<sup>20</sup> Voir les résultats de l'entreprise Pacari dans les concours internationaux : <https://www.pacari.com/nuestra-historia/>

1. La marque Pacari a reçu treize prix au nom des agriculteurs locaux qui produisent du chocolat biologique (or, argent et bronze).
2. La société équatorienne Pacari a remporté quatorze prix lors du concours international plus connu sous le nom de *Mundial del Chocolate*. C'est la marque qui a reçu le plus grand nombre de récompenses mondiales depuis le premier décembre 2014.
3. Pacari a été reconnu pour sa qualité et son innovation dans le cadre du cycle des Amériques des prestigieux *International Chocolate Awards*. La marque équatorienne a reçu vingt-deux prix : cinq en or, neuf en argent, cinq en bronze et trois prix spéciaux lors de ce concours qui s'est déroulé à New York<sup>21</sup>.
4. Le 21 novembre 2017, lors de la sixième édition des *Chocolates Ecuador Awards*, la marque Pacari reçoit neuf prix dans différentes catégories, dont la grande barre d'or avec Pacari Raw 100 % et Pacari Menta (Peralta et Barbotó)<sup>22</sup>.
5. Les plantations Encalada Villacís et l'association agricole La Fortaleza, ont obtenu le prix *International Cocoa Awards* pour l'excellente qualité de cacao produit sur son riche territoire<sup>23</sup>.
6. Dans la revue *Nouvelles d'Amérique latine et des Caraïbes* (2018), les journalistes affirment que le meilleur chocolat du monde n'est plus de Suisse ou de Belgique, le titre convoité revient à l'Équateur qui avec 18 récompenses remportées par la marque Pacari au concours *International Chocolate Awards*, le pays devient le leader du cacao aux arômes fins, une espèce connue pour ses saveurs florales et fruitées.

Ce phénomène de transformation rapide s'accompagne d'activités touristiques pendant lesquelles le visiteur peut expérimenter et participer au processus de transformation du

---

<sup>21</sup> <https://www.thisisecuador.com/blog/pacari-recibe-22-nuevos-premios/>

<sup>22</sup> <https://www.elcomercio.com/sabores/chocolate-ecuador-awards-premios-catadores.html>

<sup>23</sup> <http://ecuador.travel/trade/fr/nouvelle/1512-le-cacao-equatorien-s-est-a-nouveau-distingue-a-paris>

cacao en chocolat pour découvrir le processus de fabrication d'un produit internationalement connu tout comme pour partager l'authenticité d'une culture multiethnique attachée à son patrimoine. En moins de deux décennies, le chocolat équatorien s'est fait sa place au niveau international, mais aussi national. Le salon du chocolat qui se tient chaque année à Cumbayá à Quito est devenu une référence où les chocolatiers équatoriens exposent leur marque et leurs produits 100 % équatoriens, toute une variété de produits toujours plus innovants avec différents pourcentages de cacao de toutes les régions du pays (de 35 à 100 % de cacao), avec de nombreuses combinaisons avec d'autres produits phares de la région comme le maracuyá. Le salon offre aussi une opportunité de faire la promotion des savoir-faire de la gastronomie du chocolat, ou encore de produits innovants comme le vin rouge au chocolat<sup>24</sup>, le vinaigre, l'huile, la crème au chocolat, le vin mousseux à base de cacao, entre autres.

En conclusion, le cacao est l'un des produits les plus importants de l'histoire de l'Équateur. Il fait partie des civilisations indigènes du pays. Tout au long de l'histoire, le cacao a généré des bénéfices, même s'il est vrai qu'il y a eu des périodes de crise qui ont diminué sa demande en raison de la présence de maladies, qui ont affecté les cultures ou encore des crises sociales internationales. Cependant, la commercialisation de ce produit reste une source de travail et de subsistance pour une grande partie de la population équatorienne. Le cacao en Équateur est synonyme de culture, de développement, d'identité, de patrimoine ; c'est un héritage historique qui a participé au développement économique du pays. La richesse et l'abondance des ressources naturelles qui existent dans les zones de production du cacao ont permis le développement de saveurs et d'arômes uniques qui font partie de l'origine génétique de la fève réunissant tous les paramètres pour être un produit de référence international.

---

<sup>24</sup> <https://www.elcomercio.com/tendencias/mundo-cacao-vino-unen-salon.html>

### 1.3 Processus de transformation artisanale et industrielle

Avant de décrire chacun des processus de transformation, il est nécessaire de préciser que les fruits doivent être récoltés entre 5 à 6 mois, ce qui correspond à leur maturité optimale. Une fois que toutes les mesures de précaution ont été prises pendant le processus de maturation de la cabosse et la récolte, commence la gestion post-récolte du cacao.

Les cabosses sont cassées en évitant d'endommager et de contaminer les fèves, l'intérieur est alors placé dans un récipient afin que le mucilage soit préservé. La masse obtenue est appelée la pâte. Ce processus est connu comme écabossage (*Figure 7 : Processus de collecte des fèves*).



*Figure 7 : Processus de collecte des fèves*

*Source : Manual de cultivo de cacao para la Amazonía ecuatoriana (2009), <http://canacacao.org/wp-content/uploads/Manual-cultivo-Cacao-Ecuador.pdf>*

Une fois les fèves récupérées, commence la fermentation du cacao, un processus qui influence la qualité du grain, car il permet le développement des saveurs et des arômes caractéristiques. Il doit être réalisé dans des caisses en bois, avec des trous qui permettent la lixiviation du mucilage. Situées et protégées à l'intérieur de sacs de jute,

les fèves fermentent en général de cinq à six jours, avec un retournement de la pâte au deuxième, quatrième et cinquième jour, pour l'oxygéner et homogénéiser la fermentation (*Figure 8 : Processus de fermentation*).



*Figure 8 : Processus de fermentation*

*Source : Manual de cultivo de cacao para la Amazonía ecuatoriana (2009), <http://canacacao.org/wp-content/uploads/Manual-cultivo-Cacao-Ecuador.pdf>*



*Figure 9 : Processus de séchage*

*Source : Manual de cultivo de cacao para la Amazonía ecuatoriana (2009), <http://canacacao.org/wp-content/uploads/Manual-cultivo-Cacao-Ecuador.pdf>*

Le séchage du cacao consiste à éliminer lentement l'humidité des fèves afin d'achever la formation des arômes et des saveurs qu'exprimera le futur chocolat. La méthode de séchage dépendra des conditions des producteurs, elle peut être rudimentaire : sur le

macadam au bord de la route, ou plus élaborée comme l'illustre la *Figure 9* : Processus de séchage. Une fois les graines séchées, elles sont envoyées aux transformateurs comme matière première pour la production de chocolat, de poudre de cacao et de beurre de cacao.

La première étape du processus de transformation de la fève séchée en chocolat est la torréfaction, qui permet de modifier la couleur et la saveur, mais surtout de se débarrasser de la fine enveloppe, la coque, qui recouvre la fève. Après la torréfaction et le décorticage, les fèves sont broyées pour obtenir la pâte de cacao selon le niveau d'affinage souhaité.

### 1.3.1. Processus de transformation artisanal du cacao

Le processus de transformation du cacao en chocolat expliqué ci-dessous par étape est basé sur une révision bibliographique (Abad et coll., 2020 ; Alcívar Rodríguez et coll., 2019 ; Cadena Fiallos, 2015 ; Páez Paredes, 2015 ; Quingaísa, 2007 ; Sukha, 2016 ; Vera et Goya, 2015) et sur l'observation personnelle de la transformation du cacao de manière traditionnelle ou artisanale réalisée dans la province de Pastaza en Équateur, dans une localité qui s'appelle Porvenir (*Figure 10* : Processus artisanal : de la fève à la tablette de chocolat).

Il existe plusieurs stratégies de séchage en fonction des conditions climatiques locales. Le séchage au soleil est une des méthodes les plus utilisées par les producteurs. Les fèves de cacao sont placées sur de grandes surfaces ensoleillées comme les bords des routes. Ce processus dure environ sept jours. Les deux premiers jours doivent être réalisés en séchage contrôlé, ce qui consiste à étaler les fèves de cacao sur une épaisseur de 10 cm et à les remuer toutes les heures. À partir du troisième jour, le séchage se fera en couche mince de fèves pour atteindre sept pour cent d'humidité. Le séchage sous protection est moins courant que le séchage au soleil, cependant il devient indispensable



dans des endroits au climat plus pluvieux et chaud comme le met en évidence la *Figure 9* : Processus de séchage.

Après avoir séché, le cacao est ensuite grillé dans une grande poêle où il est mélangé lentement pour obtenir une bonne torréfaction. Le temps de torréfaction dépendra du feu et de la quantité de fèves utilisées.







*Figure 10 : Processus artisanal : de la fève à la tablette de chocolat*  
*Source : photographie personnelle*

Une fois grillé, il est étalé sur une table en bois pour procéder au décortiquage. Ensuite, les fèves de cacao sont placées petit à petit dans un moulin artisanal et sont broyées de façon à obtenir la pâte de cacao. À ce stade, l'arôme du chocolat peut être pleinement perçu. La pâte est ensuite façonnée et refroidie. Le résultat final obtenu est un chocolat artisanal 100 % cacao, fait maison sans produit ajouté (*Figure 10 : Processus artisanal : de la fève à la tablette de chocolat*).

### 1.3.2. Processus de transformation industriel du cacao

Le séchage est effectué par des méthodes modernes telles que le séchoir artificiel qui consiste en une zone de séchage avec un système de chauffage (*Figure 11 : Exemple de*

séchoir industriel). Ces améliorations industrielles présentent l'avantage de réduire le temps de séchage, d'être peu dépendants de la main-d'œuvre et de protéger l'environnement des rongeurs, des insectes, de la poussière ou de tout autre type d'agent de contamination. Les séchoirs artificiels comprennent les séchoirs à tunnel, les séchoirs solaires et les séchoirs à combustible. Les séchoirs à tunnel se composent d'un couloir couvert où se trouve une suite de plateaux dans lesquels sont disposées les fèves. La circulation d'air chaud dans ce tunnel assèche les fèves.



*Figure 11 : Exemple de séchoir industriel*

*Source : Manual de cultivo de cacao para la Amazonía ecuatoriana (2009), <http://canacacao.org/wp-content/uploads/Manual-cultivo-Cacao-Ecuador.pdf>*

D'autre part, les séchoirs solaires se composent d'un collecteur solaire où l'air est chauffé par le rayonnement, puis circule dans la chambre de séchage où sont placées les fèves de cacao. En ce qui concerne les séchoirs à combustible, ils peuvent fonctionner au bois, à la bagasse ou aux dérivés du pétrole. Le séchage se fait généralement dans des locaux fermés et par conduction.

Le cacao est ensuite torréfié, le temps et la température de torréfaction dépendent de chaque récolte, c'est-à-dire que la récolte de chaque zone est torréfiée séparément.

Avant d'être broyées, les fèves torréfiées passent dans une machine qui les décortique en triant la coque par gravité. Une fois débarrassées des derniers morceaux de coque ; trois systèmes de broyeur produisent une masse jusqu'à obtenir une pâte de cacao lisse

et soyeuse. Les fèves passent par des moulins à rouleaux qui les brisent en petits morceaux ou *nibs*, des moulins à broches qui les transforment en pâte de cacao, et des moulins à billes qui réduisent la granulométrie et la rendent lisse et soyeuse.

Ensuite, la pâte de cacao est mélangée à des additifs selon les secrets de fabrication de chaque professionnel, du sucre peut être ajouté selon le goût, un peu de beurre de cacao et de lécithine sont aussi des éléments fréquemment utilisés. Ce mélange est ensuite acheminé vers le raffineur où cinq rouleaux réduisent les particules jusqu'à une taille de 18-20 microns : à ce stade, les spécialistes peuvent déterminer les caractéristiques aromatiques du chocolat. La pâte obtenue est ensuite conchée, c'est-à-dire qu'elle est agitée à chaud d'une manière constante et soutenue pendant environ 72 heures pour améliorer son onctuosité et sa finesse selon la variété.

Le chocolat est ensuite tempéré et moulé, c'est-à-dire qu'il est refroidi sur des tables jusqu'à 28,5 degrés Celsius, puis la température est augmentée pour atteindre de 29,5 à 32 degrés Celsius selon la variété. Une fois le chocolat tempéré, il va directement dans les moules qui sont réfrigérés. À ce stade, le chocolat devient brillant, satiné, cassant et fondant. Enfin, une fois le processus de refroidissement terminé, le produit final est emballé et préparé pour être commercialisé et consommé (*Figure 12*).







*Figure 12 : Processus industriel de la fève à la tablette de chocolat*

*Source : photographie de la Finca Saquifracia - Puyo*

Ce premier chapitre nous aura permis de nous familiariser avec la réalité de la filière du cacao et du chocolat en Équateur afin de déterminer les écosystèmes linguistiques qui s'inscrivent dans l'approche de l'écologie linguistique choisie. En effet, dans une approche écologique du langage en relation avec l'étude du cacao et du chocolat en Equateur, l'étude se situe dans le domaine de la linguistique sensorielle et souligne l'importance de comprendre les aspects sensoriels de la langue dans un contexte spécifique. Elle introduit le concept d'écologie du langage en tant que cadre théorique pour étudier le langage dans son contexte socioculturel. Dans le paradigme de l'"écologie du langage", défini comme l'étude de la relation entre le langage et son

environnement physique, social et culturel, l'étude aborde l'évolution de la recherche en sciences du langage et le développement d'outils informatiques afin d'introduire l'analyse quantitative de discours authentiques. L'approche est décrite comme étant à la fois "*bottom up*" et "*top down*", les chercheurs utilisant les données quantitatives du discours pour éclairer leur analyse et leur interprétation dans un contexte écologique plus large. Ce chapitre explore également le concept de relativité linguistique, qui postule que la langue influence la manière dont les individus perçoivent et conceptualisent la réalité. Elle retrace l'évolution historique de cette idée depuis Humboldt jusqu'aux penseurs poststructuralistes plus récents. La relation entre la langue et l'environnement est abordée, en mettant en évidence des exemples tels que la langue inuit qui possède de nombreux mots pour décrire les différents types de neige et de glace. Cependant, l'hypothèse de la relativité linguistique a été remise en question par certains chercheurs et il est important de reconnaître qu'il existe un débat permanent sur l'étendue de l'influence de la langue sur la pensée et la perception. Dans l'ensemble, l'écologie linguistique et le concept de relativité linguistique en tant que cadre théorique permet de souligner l'importance de prendre en compte le contexte écologique dans lequel la langue est produite afin de comprendre ses implications sociales et culturelles. En effet, la relation entre la langue et la réalité, ainsi que des limites de l'approche linguistique générative dans la compréhension de la variation linguistique soulignent l'importance des facteurs socioculturels dans la variation linguistique et la nécessité d'approches complémentaires telles que la sociolinguistique et l'anthropologie linguistique. Le concept d'écologie linguistique, qui prend en compte l'interaction entre les langues et leurs environnements sociaux et culturels, situe le discours étudié dans son contexte de production et le positionne comme le reflet de la culture et de la construction de la réalité. La réalité équatorienne est dominée par l'importance de la diversité linguistique et l'impact de la domination linguistique d'une culture coloniale sur les langues et les cultures indigènes. Mais ce n'est pas le seul aspect, la biodiversité naturelle de l'Equateur alimente naturellement le lien entre la diversité linguistique et la biodiversité environnementale. Cette réalité explique la diversité des écosystèmes linguistiques indissociables de l'étude du comportement des langues et des discours

produits. Dans ce chapitre, il a été considéré important de mentionner au lecteur les défis auxquels les communautés autochtones sont confrontées en raison de la colonisation et de la perte des langues ancestrales ; ainsi que l'impact des facteurs sociaux et culturels sur les relations de pouvoir et de prestige qui influencent la vitalité d'usage d'une langue. La dynamique du pouvoir et les pressions entre les différents groupes humains en tant que facteurs clés du changement et du déplacement des langues, les politiques linguistiques et le rôle du pouvoir politique dans la détermination du statut et de la préservation des langues sont des thématiques qui permettent de comprendre les mécanismes de substitution et à d'abandon des langues en Equateur ; ainsi que l'émergence de l'activisme linguistique et de l'écolinguistique politique pour la préservation de la diversité linguistique. Cette réalité sociale est la trame de fond des conflits qui existent entre les différents écosystèmes linguistiques. Le contexte social et la dynamique du pouvoir au sein de chaque écosystème, influencée par des variables socioculturelles telles que la profession, le sexe et le statut social, apparaissent comme un savoir préalable indispensable à l'interprétation de l'analyse de discours pour comprendre la régulation et l'expression de la langue.

Les modes de communication et les discours utilisés par les professionnels impliqués dans la culture du cacao et la production de chocolat en Equateur sont construits dans cette réalité linguistique conflictuelle. Les différences d'utilisation de la langue, les intentions de communication, ainsi que le contexte écologique et culturel influencent la production langagière des professionnels de l'industrie du cacao et du chocolat. Le rôle de la perception et de l'environnement dans la formation des processus cognitifs et des expressions linguistiques liés aux expériences de dégustation du chocolat est aussi modulé par ces paramètres socioculturels : le cacao est considéré comme un symbole de fierté nationale.

La représentation sociale du cacao et du chocolat en Equateur, ainsi que l'organisation du système de production et de fabrication sont aussi des savoirs préalables indispensables à l'interprétation des discours étudiés. Le processus de production du

chocolat est segmenté : soit les agriculteurs vendent les fèves de cacao séchées à des professionnels de la transformation, soit ils les vendent à des exportateurs.

De façon à compléter la description de l'écosystème linguistique, le chapitre offre au lecteur une description détaillée de la fève de cacao, de son processus de culture et de transformation ainsi qu'un bref historique du cacao en Équateur, de son acceptation à l'international, et de l'importance culturelle et économique comme symbole de l'identité et du patrimoine du pays.

Dans le cadre théorique de l'écologie linguistique, ses informations donnent au lecteur qui n'est pas familiarisé avec la réalité de l'industrie du cacao équatorien un aperçu indispensable des paramètres socioculturels qui modulent la production langagière des professionnels du domaine.

## Chapitre 2 : Le corpus en linguistique sensorielle

---

Une des priorités communes à de nombreux domaines professionnels est de communiquer les vertus d'un produit ou d'un service de façon à motiver son acquisition tout en suscitant le plaisir par l'expérience sensorielle, par l'émotion, et s'assurer de la pleine satisfaction du client. Le domaine du cacao et du chocolat est pleinement exposé à cette équation en raison des caractéristiques intrinsèques du produit qui font appel aux cinq sens pour satisfaire l'acquéreur. Par conséquent, dans le monde du cacao et du chocolat, la communication est étroitement liée au sensoriel. Ainsi, l'analyse discursive de situations communicationnelles produites dans l'environnement équatorien de la production du cacao et du chocolat s'inscrit parfaitement dans un cadre de la linguistique sensorielle qu'il est nécessaire de définir et illustrer avant de présenter la typologie de la terminologie sensorielle considérée dans cette étude.

Il n'existe que peu de place pour l'émotion lorsque les résultats d'un examen approfondi de la composition organoleptique d'une tablette de chocolat sont présentés sous la forme d'un tableau de noms de molécules et de leurs différents pourcentages. Peu nombreuses sont les stratégies de promotion qui mettent en avant un discours d'ingénieur chimiste. Cependant, dans un domaine où le sensoriel est un critère de valorisation important pour l'évaluation d'un produit, les professionnels de l'industrie du chocolat sont confrontés à l'utilisation incontournable d'un discours internationalement normé qui s'inspire des pratiques langagières de l'industrie viticole pour décrire et comparer leurs produits. Cette stratégie au service d'une analyse comparative des qualités d'un produit fini se limite le plus souvent à un glossaire terminologique de descripteurs évaluatifs. La maîtrise professionnelle développée par les experts pour sélectionner le descripteur qui décrit le mieux la réaction psychophysique induite par le chocolat reste dépendante d'un jugement perceptif, exercé par des experts certes, mais ces experts n'en demeurent pas moins des acteurs sociaux ordinaires qui utilisent par ailleurs des formes de langage



ordinaires. Même si le choix des descripteurs est large, il arrive que les experts soient contraints de combler un vide dénominatif, ce qui peut les amener à développer un nouveau descripteur pour renouveler leur discours. Le choix des stratégies langagières produites en situation de communication pour exprimer au mieux la perception et l'interprétation d'une sensation offre aux chercheurs un champ d'étude récent et prometteur pour rendre compte de l'expérience sensorielle. L'étude de ces stratégies qui vont de la terminologie chimique aux appréciations et aux marqueurs hédoniques, en passant par la taxonomie des descripteurs, est un des objectifs de la linguistique sensorielle. Résultat d'une évolution logique de la diversification de la recherche linguistique poststructuraliste, elle se situe au carrefour des neurosciences cognitives et de la linguistique. Dans le cadre d'une linguistique contextuelle centrée sur l'analyse du discours sensoriel, elle utilise l'analyse de corpus compilés et les outils de traitement automatisé du langage.

Contrairement aux approches basées sur la terminologie du sensoriel définis a priori, le concept de sensoriel se construit à partir des cinq sens, sur l'expérience sensitive du moment et a posteriori sur le souvenir stockée dans la mémoire à long terme en incluant les sensations multimodales influencées par l'expérience. La nature même de ces définitions place l'écosystème socioculturel d'une communauté linguistique dans la dynamique de la production discursive qui sous-tend la manière dont un individu exprime sa perception et son interprétation de la réalité. Le sensoriel se nourrit des connaissances acquises par les sens, et sa restitution discursive met en scène l'espace spatio-temporel délimité par l'écosystème sociolinguistique de manière résolument pragmatique. C'est ainsi qu'il apparaît que le discours se construit à partir de la réalité culturelle en plus des traits précis des variables socioprofessionnelles déterminées par les différentes communautés discursives dans lesquelles l'acteur social évolue au cours de la production langagière. Il apparaît également que le discours n'est pas seulement ajusté à l'intention de communication de l'acteur social. L'intention de communication est définie dans cette proposition comme étant d'une part, l'expression de la sensorialité gustative et olfactive liée au cacao et au chocolat, qu'elle soit suscitée ou vécue, et

d'autre part, l'expression de la perception cognitive du cacao ou du chocolat que possède l'acteur social.

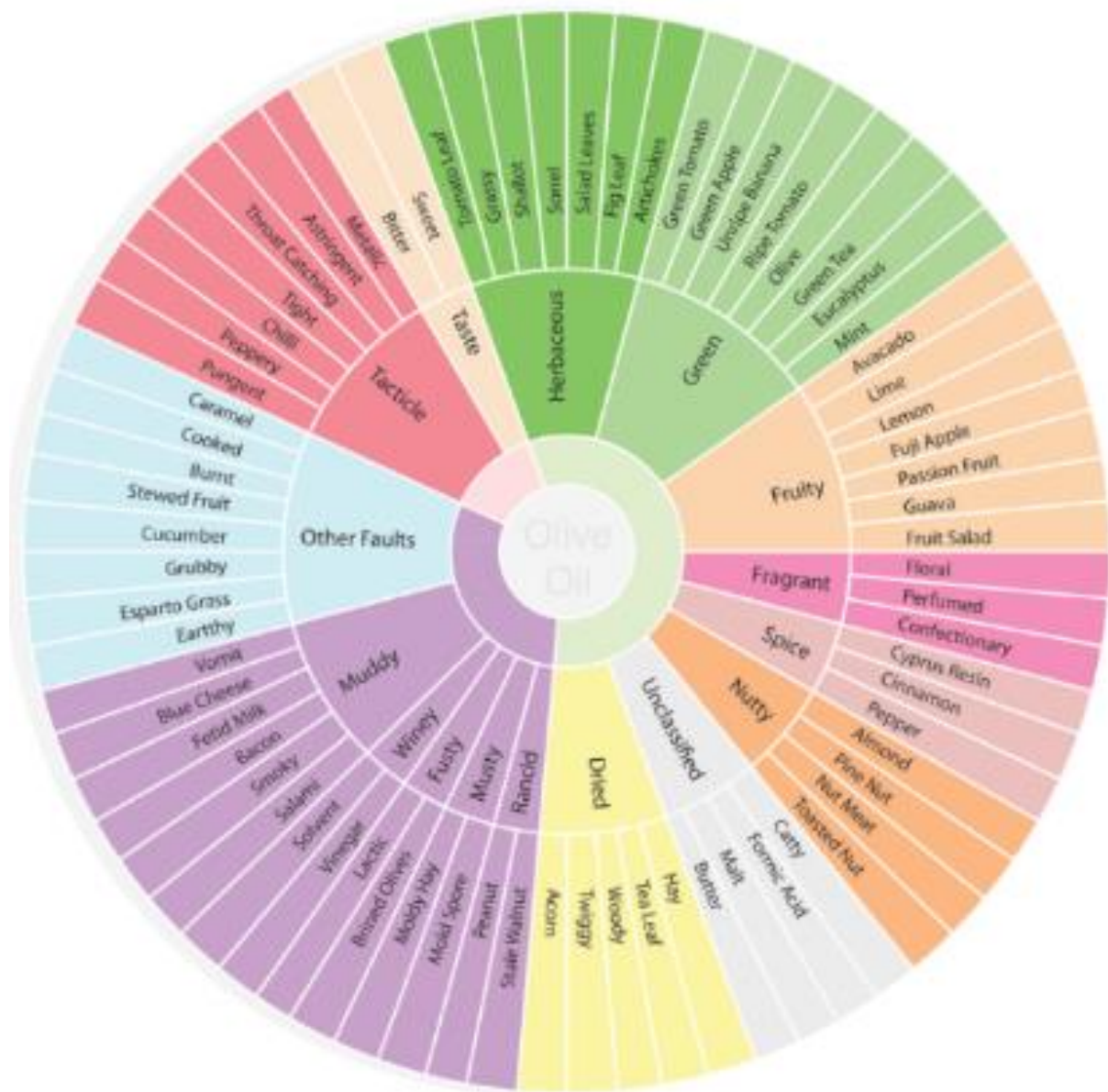
Cette pertinence discursive émerge, se développe et s'épanouit dans une réalité encodée dans un écosystème linguistique particulier qui permet le développement d'hypothèses anticipatives. Par conséquent, le discours généré par des situations de communication spécialisées est pris en compte dans cette étude à partir de la compilation d'un corpus qui est à la fois représentatif de la situation de communication réelle et de l'environnement dans lequel il est produit. Pour partager cette interprétation avec son public, l'acteur social, qu'il soit professionnel ou *aficionado*, a l'obligation de créer un discours qui n'est rien d'autre que la matérialisation physique de la représentation sociale que l'individu se fait de son interaction avec un objet particulier, en l'occurrence le chocolat. L'objectif de cette étude étant de décrire une partie du processus de matérialisation, la compilation d'un corpus représentatif devient le sujet d'étude afin de décoder le processus sémantique de production langagière en espérant avoir accès aux intentions de communication à partir de l'analyse des différentes matérialisations physiques à l'aide d'une approche située et outillée. Les inférences proposées à partir des résultats d'analyse des corpus et sous-corpus représentatifs de l'expression qu'un acteur social d'une même communauté langagière peut avoir de la réalité au travers d'une émotion stimulée ou d'une perception cognitive, permettent de modaliser les stratégies discursives utilisées pour partager le sensoriel tout en laissant un espace suffisant de sérendipité propre à cette approche « bottom-up/top-down » (Brown et Yule, 1983: 234-235) pour souligner de possibles résultats inattendus.

## 2.1 Positionnement théorique de la recherche : Linguistique sensorielle et écologie du langage

Cette section propose de montrer comment la présente démarche réflexive peut être compatible avec certains des différents courants de pensée qui ont alimenté et continuent d'alimenter les recherches en sciences du langage. La chimie offre tout un

lexique créé à partir de nombreuses associations d'affixes pour nommer une molécule. Le résultat d'une analyse précise de la composition organoleptique d'une barre de chocolat reproduit sous forme d'un tableau de noms de molécule avec leurs différents pourcentages ne laisse que peu de place à l'émotivité. Il est d'ailleurs rare qu'un professionnel en marketing utilise cette stratégie pour vanter les vertus du chocolat. Cependant, les professionnels de la filière du chocolat immergés dans le monde du sensoriel éprouvent le besoin de caractériser leurs produits, que ce soit par des moyens extralinguistiques ou le plus souvent par le biais d'une production langagière. Ils doivent choisir un discours et se mettre d'accord sur le lexique ou les stratégies discursives à utiliser tout en faisant face aux différentes réactions psychophysiques, les sensations, que chaque composant peut stimuler chez un individu (Dubois, 2010). Actuellement, l'identification et la description des propriétés sensorielles a tendance à créer une normalisation consensuelle du langage pour définir des descripteurs lexicaux dont le but est d'assurer l'interprétation et la dénomination d'une sensation à partir d'une taxonomie contrôlée, aussi connue comme la carte d'identité linguistique des odeurs (Kleiber, 2012) ou encore plus fréquemment pour le goût comme roues des arômes qui le plus souvent se limitent à un glossaire terminologique de descripteurs évaluatifs (*Figure 13* : Roues des arômes du chocolat, de l'huile d'olive et du vin).





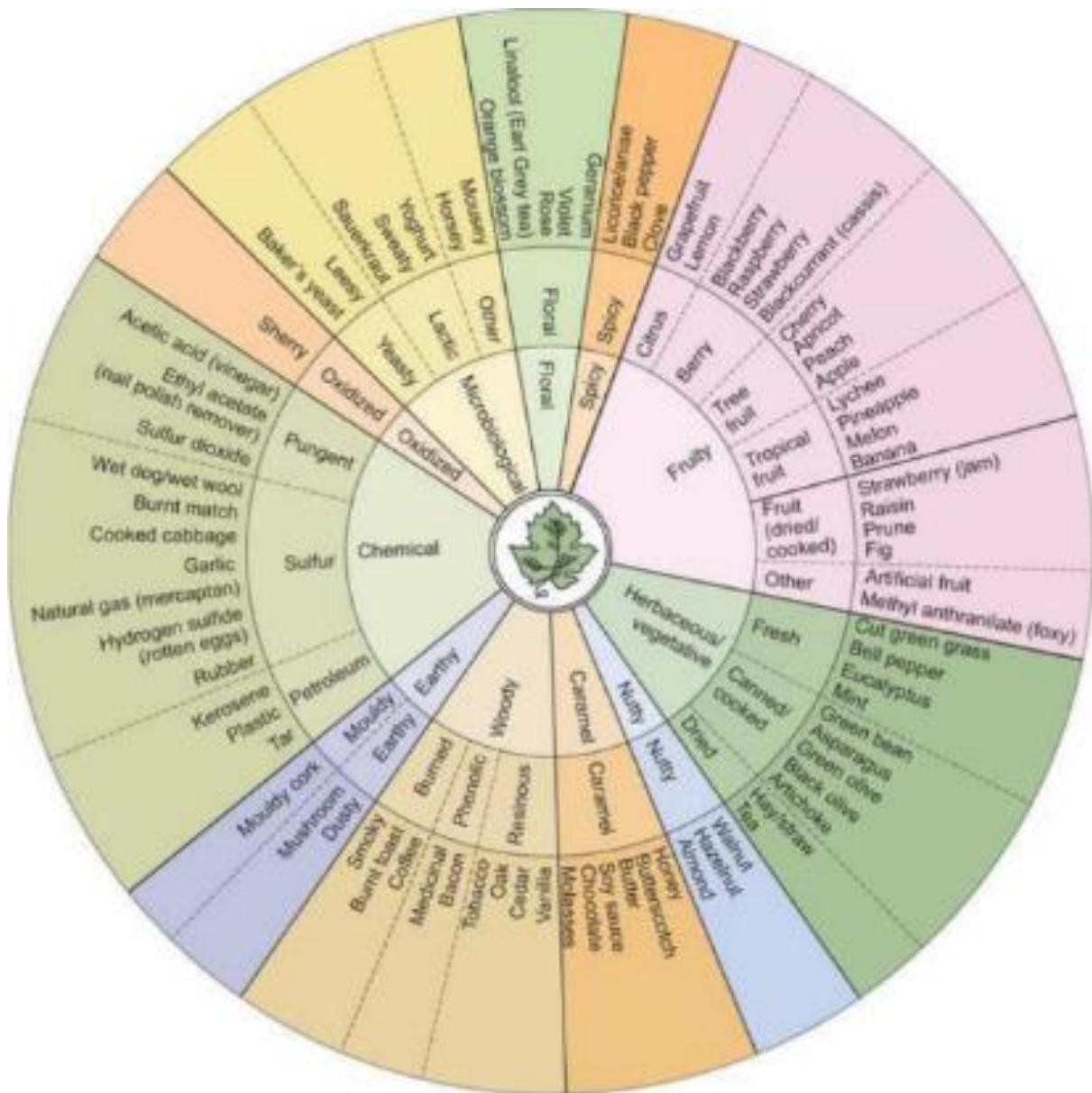


Figure 13 : Roues des arômes du chocolat, de l'huile d'olive et du vin<sup>25</sup>  
 Source : élaboré par Sukha (2016)

La maîtrise professionnelle développée par les experts pour choisir le descripteur qui décrit au mieux la réaction psychophysique stimulée par le chocolat reste tributaire d'un jugement perceptif, exercé par des experts certes, mais ces experts n'en demeurent pas

<sup>25</sup> [https://www.cocoaqualitystandards.org/fileadmin/templates/CocoaQuality/Uploads/Documents-and-reports/REVIEW-Cocoa\\_Quality\\_Flavour\\_Standards\\_-\\_Darin\\_SUKHA\\_sans\\_Appendices-12May2016-ESPANOL.pdf](https://www.cocoaqualitystandards.org/fileadmin/templates/CocoaQuality/Uploads/Documents-and-reports/REVIEW-Cocoa_Quality_Flavour_Standards_-_Darin_SUKHA_sans_Appendices-12May2016-ESPANOL.pdf)

moins « des sujets ordinaires, qui de surcroît utilisent des formes (y compris lexicales) du langage ordinaire » (Morange, 2009 : 139). De plus, aussi riche que puisse être l'éventail des descripteurs, certaines situations mettent les experts face à un vide dénomiatif (Kleiber, 2012) et les obligent à entrer dans un processus créatif qui peut aboutir à l'émergence d'un nouveau descripteur. Jusqu'ici, nous avons considéré la production langagière de l'expert qui possède un savoir lexical pragmatique étendu et actualisé pour combler au mieux cette absence de mot ; savoir auquel le simple consommateur, ou même l'aficionado, n'a pas toujours accès ce qui ne l'empêche pas pour autant de s'exprimer.

D'une terminologie chimique, aux appréciations et marqueurs hédoniques, en passant par la taxonomie des descripteurs, le choix des stratégies langagières produites dans les situations de communications pour exprimer au mieux la perception et l'interprétation d'une sensation offre aux chercheurs un champ d'étude récent et prometteur pour rendre compte de l'expérience sensorielle. L'étude de ces stratégies est une des raisons d'être de la linguistique sensorielle. À la croisée des neurosciences cognitives et de la linguistique, la linguistique sensorielle est le résultat d'une évolution logique de la diversification des recherches en linguistiques poststructuralistes. Inspirée de la linguistique cognitive (Gardner, 1987). Elle utilise les outils de la linguistique de corpus (Fillmore, 1992 ; Sinclair, 1992 ; Biber et coll., 1998 ; Rastier, 2004 ; Baker, 2016 ; Stefanowitsch, 2020) et du traitement automatisé des langues (Eisenstein, 2019, Jurafsky et Martin, 2019) dans le cadre d'une linguistique contextuellement située (Fairclough, 2005 ; Longhi, 2009, 2018 ; Condamines et Narcy-Combes, 2015 ; Méric, 2016 ; Méric et Gautier, 2017 ; Sharifian, 2017) et spécialisée dans l'analyse des discours du sensoriel (Digonnet, 2018, Gautier et Bach, 2017 ; Bach, 2020a ; Gautier, 2019 ; Méric, 2019 ; Cahuana, 2021).

La recherche autant sur *Google Scholar*© en français que sur d'autre moteur de recherche comme *Semantic Scholar*© de l'expression « discours du sensoriel » ne



retourne pas plus d'une trentaine de résultats<sup>26</sup> majoritairement liés aux travaux de recherches de l'équipe dijonnaise du laboratoire TIL (Texte, Image, Langage) réunie par le professeur Gautier dont les membres sont entre autres Parizot (2019a, b), Mancebo (2019a, b), Cahuana Velasteguí et coll. (2021a, 2021 b), Bach (2020a, 2020 b, 2021a) et Méric (2016, 2019, 2021). Plusieurs publications des membres de l'équipe en relation au vin, au crémant, ou encore au chocolat, font référence aux notions de discours et de sensoriel. En ce qui concerne le concept de discours, inspiré de la définition proposée par Bach : « molécule dynamique et cognitive composée d'atomes plurisémiotiques et conceptuellement liée à l'individu, à une communauté spécifique et à une réalité sociale »<sup>27</sup> (2020a : 2) ; nous retiendrons pour notre étude une définition liée au concept d'écologie du langage précédemment considéré. Le discours peut être alors défini comme la matérialisation plurisémiotique, dynamique et cognitive d'un message situé et immergé dans un écosystème socioculturel construit selon l'architecture conceptuelle individuelle et collective qu'un membre d'une communauté langagière a de la réalité. Quant au concept de sensoriel, nous le considérerons basé sur les cinq sens et reposant sur l'expérience actuelle et a posteriori stockée en mémoire à long terme en incluant les sensations multimodales influencées par l'expérience. Cette conceptualisation du sensoriel s'oppose aux approches sur les terminologies sensorielles définis à priori.

Le contenu même de ces définitions positionne l'écosystème socioculturel d'une communauté langagière au sein de la dynamique discursive qui rend compte de la perception et de l'interprétation de la réalité exprimée par un individu dans le cadre de l'écologie linguistique. Dans le cas du discours sensoriel associé au monde du chocolat, cette expression concerne les situations de communication spécifiques aux pratiques socioprofessionnelles de tout un domaine d'activité dans lequel le plurisémiotisme précédemment souligné évolue selon trois axes : l'axe socioculturel, l'axe du domaine de spécialités, et l'axe de perception et de conceptualisation individuelle. Le sensoriel

---

<sup>26</sup> Recherche effectuée en juin 2023

<sup>27</sup> Traduit par nous de : “*dynamic cognitive molecules, composed of plurisemiotic atoms, conceptually linking an individual, a specific community and a social reality.*”



se nourrit de l'expérience acquise au travers des sens, et sa restitution discursive est résolument pragmatique mettant en scène l'espace spatio-temporel défini par les trois axes précédents. La conception triadique de ce plurisémiotisme, de cette approche du discours du sensoriel, offre un cadre théorique sémiotique peircien spécifique adapté à la pragmatique de la linguistique sensorielle située dans son écosystème. En accord avec Everaert-Desmedt qui considère que le « processus sémiotique selon Peirce intègre toutes les composantes de la sémiotique : la pragmatique (domaine de l'interprétant) est indissociable de la sémantique (domaine de l'objet) et de la syntaxe (domaine du *representamen*) » (2011 : sn). Il est à préciser que dans la sémiotique Peircienne, le *representamen* est directement lié au concept de signe qui représente l'objet. Le contexte socioculturel équatorien dans lequel se produit le discours d'étude offre une grande amplitude sémantique étroitement liée à l'environnement direct du locuteur qui exprime les représentations personnelles de ses émotions au travers d'une syntaxe textuelle et d'un lexique qui lui est propre.

Ainsi, dans le cadre spécifique de cette analyse, nous proposons que le discours du sensoriel, objet d'étude de la linguistique sensorielle, soit considéré comme la trace physique d'une production langagière qui communique la perception d'une sensation. L'acteur social transmet cette expérience sensorielle à une audience pour que les membres de cette audience puissent en interpréter le sens à partir de l'environnement socioprofessionnel et socioculturel de la communauté langagière à laquelle ils appartiennent, de leur propre représentation sociale et du domaine de spécialité. Le discours est non seulement ajusté à l'intention de communication de l'acteur social, mais il est aussi construit à partir de la réalité culturelle qui, complétée par les caractéristiques mêmes des variables socioprofessionnelles, détermine les différentes communautés discursives dans lesquelles l'acteur social évolue lors de sa production langagière. Ainsi, l'épaisseur du discours produit dépend des caractéristiques sociales, professionnelles et culturelles du locuteur et de son intention de communication ; en résumé de l'écosystème dans lequel la production langagière est située. Dans cette proposition, l'intention de communication est fixée comme étant d'une part

l'expression de la sensorialité liée au cacao et au chocolat, que cette sensorialité gustato-olfactive soit évoquée ou expérimentée, et d'autre part l'expression de la perception cognitive que l'acteur social a du cacao ou du chocolat.

Dans la conceptualisation triadique de la sémiotique peircienne, le discours de la sensorialité est donc tributaire de l'émotion que provoque la sensation - évoquée ou expérimentée - stimulée par l'objet, de l'interprétation issue de la perception cognitive du *representamen*, et de la modélisation qui matérialise la représentation sociale de l'objet. La pertinence du discours sensoriel produit dépend de la sensorialité, de l'écosystème culturel, et de la perception de la réalité que l'acteur social a quand il matérialise par sa production langagière ses émotions et ses interprétations. Cette pertinence discursive naît, croît et s'épanouit dans une réalité inscrite dans un écosystème linguistique spécifique qui permet la création d'hypothèses anticipatoires (Reboul, 1992). Il en résulte que cette étude prendra en considération le discours produit à partir de situations de communication spécialisées compilées dans un corpus représentatif aussi bien de la situation de communication authentique que de l'écosystème dans lequel elle est produite. Ce corpus est donc considéré comme la matérialisation représentative de l'expression qu'un acteur social d'une même communauté langagière peut avoir de la réalité au travers d'une émotion stimulée ou d'une perception cognitive. La *Figure 14* : Écologie linguistique de la production langagière en linguistique sensorielle illustre le processus de sémiotisation prévio à la production langagière de l'interprétant qui une fois compilée constitue le corpus d'étude.

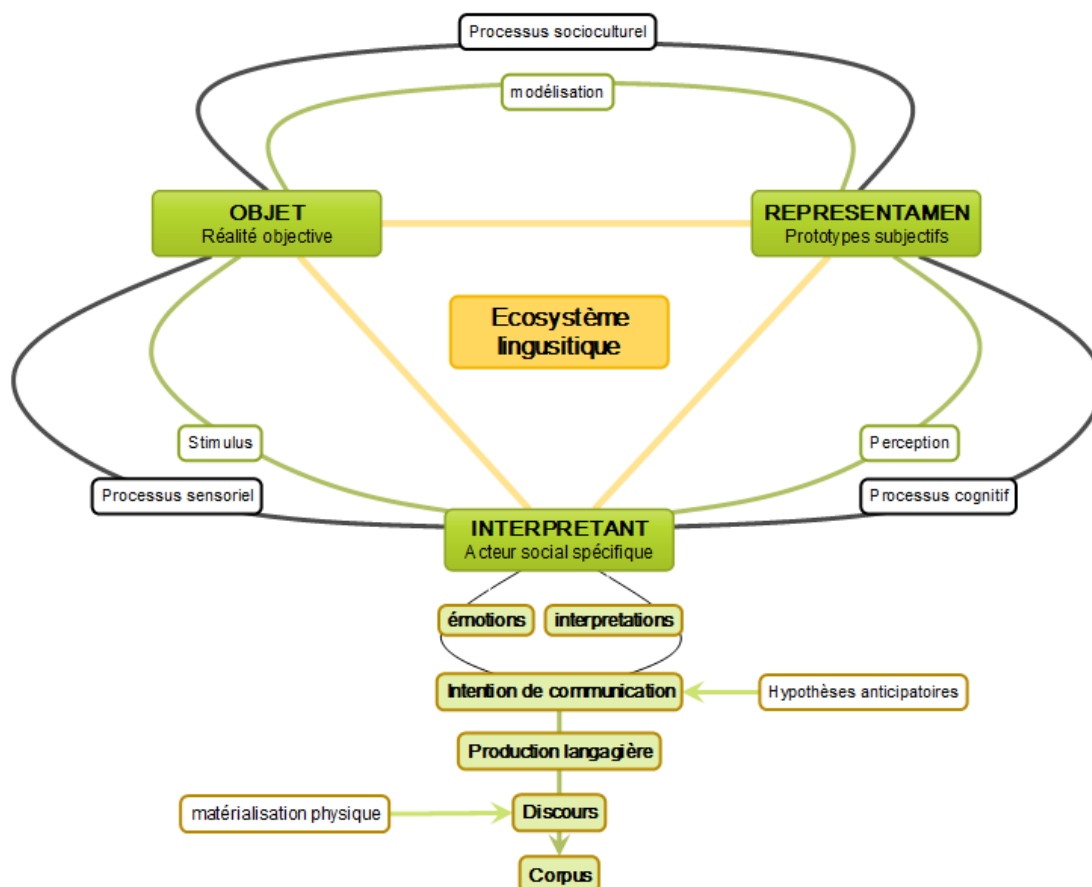


Figure 14 : Écologie linguistique de la production langagière en linguistique sensorielle

Source : élaboré par nous

En raison de la nature même de l'objet d'étude, selon un processus propre à la linguistique de corpus, ce sont les outils liés à la textométrie et à l'analyse du discours qui sont utilisés pour offrir au chercheur des résultats quantitatifs dont l'analyse met en évidence les implications que les caractéristiques prototypiques des interprétants ont sur le discours de la sensorialité. Ces outils s'appliquent dans le cadre d'une recherche exploratoire et descriptive de différentes stratégies de construction du sens. Ils sont appliqués pour exprimer une sensation au travers de matérialisations linguistiques compilées à partir de situations authentiques de communication situées selon un raisonnement abductif dont les inférences permettent « de former des conjectures à partir d'indices et de traces » (Angué 2009 : 79) tout en laissant un espace suffisant de

sérendipité propre à cette approche « bottom-up/top-down » (Brown et Yule, 1983 : 234–235) pour se laisser surprendre par les résultats obtenus. Dans la conceptualisation peircienne, l'abduction se présente comme une méthode d'interprétation reliant le « sémiotique » au « pragmatique » en se fondant sur l'exploration et la capacité à révéler l'inattendu. La théorie de l'inférence abductive de Peirce commence par la perception d'un signe ou l'observation d'un phénomène, quel que soit sa nature, et termine par la supposition d'un principe général qui expliquerait la réalité perçue ou observée introduisant la notion d'inférences plausibles autrement appelées interprétations. Pour compléter le raisonnement, il est nécessaire de faire appel aux inférences inductives et déductives qui assumeront les rôles de justification et validation des inférences abductives. Pour illustrer ce propos, il est important de comprendre que le *representamen* chocolat noir, par exemple, n'a pas de signification inhérente. Le signe noir n'a de sens que lorsqu'il est interprété par la compréhension commune du sens du mot « noir » généralement partagée par les membres d'un même écosystème sociolinguistique. Le lien entre les lettres n et o, ou le lien entre la combinaison de lettres du mot noir et la couleur noire, est donc une coutume sociale et une norme linguistique. Cependant, si le même signe - le mot noir - était présenté avec l'objet chocolat, il évoquerait probablement dans l'esprit de celui qui le voit des représentations sociales qu'il sera intéressant de découvrir à partir de l'analyse abductive des corpus proposés.

## 2.2 Typologie de la terminologie sensorielle selon trois grands ensembles

La diversité discursive des différents intervenants de la chaîne de production du chocolat en Équateur est aussi riche et variée que l'environnement socioculturel dans lequel ils évoluent. En effet, il y a autant de différences entre la barre de chocolat et la fève de cacao qu'entre le producteur de chocolat et l'agriculteur qui cultive cette même fève. Chaque acteur social impliqué vit et exprime ses émotions dans sa propre réalité

et chaque émotion est perçue dans un contexte propre à chaque réalité. Que cette diversité soit multi-, inter-, pluri-, ou encore transculturelle, selon l'hypothèse de Sapir-Whorf, elle interfère à tous les niveaux du processus sémiotique aussi bien lors du processus sensoriel, à l'origine de l'émotion, ou du processus cognitif, à l'origine de la perception, que lors de l'interprétation sémantique qui précède l'expression linguistique de l'intention de communication. L'émotion issue de l'expérience sensorielle est un objet dynamique, elle naît soit d'un stimulus expérimenté dans le cas d'une dégustation par exemple, soit d'un stimulus évoqué. La perception est issue d'un processus de cognition, quand un acteur social réfléchit pour répondre à une question du type : « Si je vous dis : chocolat, à quoi pensez-vous ? » Émotion et perception se trouvent être à l'origine de l'interprétation de la réalité sociale du locuteur laquelle fait naître chez tout être social une intention de communication. Il éprouve le besoin de produire un discours pour partager cette interprétation avec son audience. Ce discours n'est autre que la matérialisation physique de la représentation sociale personnelle de l'expérience qu'un acteur social a avec un objet qui dans notre cas n'est autre que le chocolat. Cette matérialisation compilée dans un corpus représentatif devient le sujet d'étude de cette recherche dans le but de décrire une partie du processus d'interprétation de l'acteur social. La compilation de ces matérialisations, produites dans les situations de communication spécifiques étudiées, constitue les corpus et sous-corpus d'analyse. L'interprétation n'est autre que le résultat du processus sémantique effectué par un acteur social qui appartient à une communauté de langue représentative des différents discours compilés. La recherche proposée prétend décoder le processus sémantique de production des interprétations à partir de l'analyse des différentes matérialisations physiques que laisse l'acteur social dans son écosystème discursif (*Figure 14 : Écologie linguistique de la production langagière en linguistique sensorielle*) à l'aide d'une approche située et outillée. Nous avons identifié et considéré trois groupes de descripteurs dont les caractéristiques seront décrites en détail dans l'analyse textométrique :

a. - Descripteurs objectivistes de typologie du cacao, ils représentent une terminologie normée d'un discours académique et professionnel dont l'intention de communication est la caractérisation et l'évaluation du produit pour une audience le plus souvent internationale. Les occurrences de ce type de descripteur sont le plus souvent élevées dans des discours impliquant des nomenclatures, des roues des arômes, avec comme objectif la création de références ou de critères de comparaison pour une évaluation quantitative normée du produit.

b. - Descripteurs sensoriels, ils sont tributaires de facteurs extralinguistiques situés faisant appel à l'expérience et aux connaissances propres du locuteur qui les utilise dans un contexte spécifique. Nous pouvons, d'ores et déjà, considérer l'exemple suivant : *pepa de oro* qui fait directement référence avec le patrimoine culturel sud-américain et la sensation de perfection à travers la représentation sociale de l'or en Équateur.

c. - Descripteurs hédoniques, souvent associés à l'expression d'une expérience sensorielle placentaire ou désagréable. Positifs, ils font référence au plaisir des sens, à la satisfaction ; négatifs ils se réfèrent à un dégoût. Ils permettent l'évaluation qualitative personnelle des sensations en relation directe avec le contexte discursif au moment précis de la production langagière.

Après avoir décrit aussi bien le cadre théorique général que particulier de cette étude, il convient de s'intéresser au processus de compilation du corpus et à la description des situations de communication qu'il représente.

## 2.3 Processus de compilation du corpus d'étude

Dans son acception linguistique, le corpus est considéré par le TLFi (Trésor de la langue française informatisé) comme un « ensemble de textes établis selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique <sup>28</sup> ». Même si le TLFi adhère au principe selon lequel un mot n'a de sens et de valeur que dans son contexte ; même s'il propose des définitions d'usage en citant des exemples que l'on trouve dans la littérature française, cette définition est plus axée sur la fonction du corpus que sur sa nature même, et semble optimiste dans son critère d'exhaustivité. Il est en effet improbable que l'on puisse rassembler ou compiler tous les textes qui relèvent d'un thème ou d'une situation de communication spécifique.

Même si les termes critère thématique et exemplaire ne définissent pas clairement la nature d'un corpus, ils suggèrent la nécessité d'une certaine représentativité de la langue, ce qui semble utopique si le but est de représenter une langue et son usage par le biais d'un échantillonnage, aucun des deux termes n'explique vraiment comment cet ensemble de textes a été créé. Cependant, cette définition a l'avantage de mettre en avant le "texte", dont Pery-Woodley (1995 : 221) souligne également qu'il est une composante essentielle du corpus : « Un corpus se compose par définition de discours, de langue "concrète" [...], et c'est inmanquablement sous la forme de textes – écrits ou parlés – que la langue se réalise en discours ». Même si elle introduit la notion de texte, la définition proposée par le TLFi est ambiguë quant à la nature de l'ensemble des textes et de leur représentativité. La proposition de Habert (2000 : 11) semble plus précise : « Un corpus est une collection de données langagières qui sont sélectionnées et organisées selon des critères linguistiques et extralinguistiques explicites pour servir d'échantillon d'emplois déterminés d'une langue. »

L'auteur met en évidence les concepts de sélection et d'organisation selon des *critères explicites* pour représenter au mieux un phénomène déterminé ; mais il remplace l'idée

---

<sup>28</sup> <http://www.cnrtl.fr/definition/corpus/substantif>

de texte par celle de données linguistiques. Par conséquent, il est nécessaire de rechercher une définition plus approfondie tout en conservant le concept fondamental de texte, en particulier en ce qui concerne les critères explicites. La conception praxéologique de Rastier semble être pertinente avec la position théorique de cette étude. Rastier soutient que le discours ne peut pas être séparé du contexte dans lequel il est construit et interprété définissant le corpus de la manière suivante :

Un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d'une gamme d'applications. Rastier (2005, 32).

La définition de Rastier affirme que le corpus est et restera un ensemble de textes organisés et prêts à être analysés, mais il introduit également l'idée d'intégralité, qu'il applique aux éléments constitutifs du corpus et élimine la possibilité qu'un fragment de texte y soit inclus.

Rastier affirme que la structure interne du corpus doit tenir compte des discours et des genres, être représentative d'une situation de communication choisie et présenter une certaine homogénéité pour assurer la spécification et la caractérisation de ses éléments constitutifs. Dans ce contexte, il semble essentiel de disposer de normes permettant de déterminer si un texte fait partie ou non du corpus qui n'est pas un outil de travail, mais plutôt un objet d'étude, structuré et organisé pour représenter les phénomènes linguistiques associés à une situation de communication particulière. La dernière section de la définition de Rastier rappelle l'affirmation de Teubert (1996 : 240) selon laquelle « les corpus sont les ressources les plus appropriées pour les études empiriques sur le langage », ce qui semble particulièrement adapté à l'étude du discours sensoriel.

La représentativité du corpus compilé dépend de la description détaillée de la situation de communication dans laquelle il est produit. Bider, Connor et Upton considèrent qu'un corpus construit autour d'une approche *bottom-up* « doit fournir une description linguistique intelligible [des] caractéristiques généralisables de l'organisation



discursive commune aux textes du corpus étudié<sup>29</sup> » (2007 : 156). Comme objet de recherche, le corpus doit être compilé pour être la représentation fidèle des phénomènes linguistiques associés à une praxis sociale spécifique culturellement ancrée dans son écosystème de production langagière.

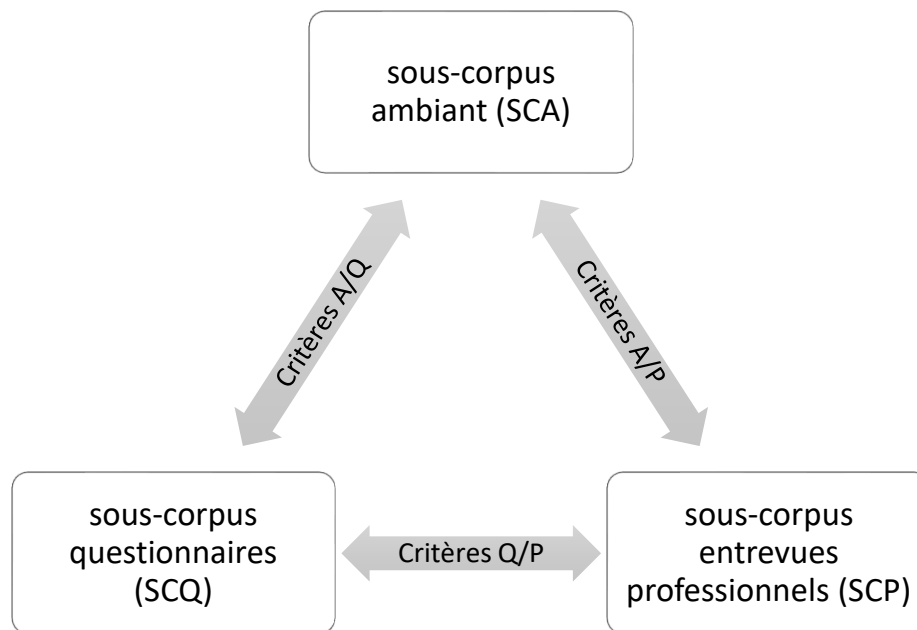
Pour l'étude du processus de sémantisation des sensations, et des émotions qui en résultent, trois sous-corpus ont été considérés et positionnés dans le *tertium comparationis* de la *Figure 15* : Paramètre du *tertium comparationis* constituant ainsi une comparaison triadique des critères communs aux trois sous-corpus qui présentent une variation entre deux d'entre eux. Selon le *Tableau 1*<sup>30</sup>, les critères dont les variations peuvent exercer une influence directe sur la production langagière entre les sous-corpus ambiant et questionnaires sont : Contexte social, Interprétant, Nombre, Présence, Type d'interaction, Auteur, Énonciateur, Factualité, Fonction, Thème, Situation géographique, période, Tokens. De même, lors de l'analyse, il est important de considérer entre les sous-corpus ambiant et entretiens professionnels les variations des critères suivants : Canal de production, Canal d'énonciation, Support d'analyse, Temporalité, Contexte social, Interprétant, Nombre, Présence, Type d'interaction, Auteur, Énonciateur, Factualité, Fonction, Thème, Situation géographique, Période, Tokens. De façon à clore la comparaison triadique, l'interprétation des résultats d'analyse prend aussi en compte, entre les sous-corpus questionnaires et entretiens professionnels, les variations des critères suivants : Canal de production, Canal d'énonciation, Support d'analyse, Temporalité, Contexte social, Auteur, Énonciateur, Factualité, Fonction, Thème, Situation géographique, Période, Tokens. De façon à faciliter la compréhension, il est important de définir le *token* comme une unité de discours qui peut présenter des normes différentes. Dans le programme Orange, les

---

<sup>29</sup> Traduit par nous : “*should provide a comprehensive linguistic description [of] generalizable patterns of discourse organization that hold across all texts of the target corpus*”

<sup>30</sup> Le tableau présenté est une adaptation du tableau utilisé par Méric (2017 : 447) à partir duquel les critères ont été modifiés pour que la description des paramètres des situations de communications soit pertinente avec l'écosystème linguistique de l'étude.

*tokens* ont été définis comme des unités en excluant tous les numéros et les signes de ponctuation.



*Figure 15 : Paramètre du tertium comparationis*

*Source : élaboré par nous*

Les variations des critères considérés sont identifiées dans le *Tableau 1*, ainsi chaque écosystème sociolinguistique, dans lequel les discours ont été produits, est caractérisé aussi bien par ses similitudes que ses différences. Parmi ces critères, deux d'entre eux ont été déterminants pour la sélection des sous-corpus : l'expertise du locuteur et l'ampleur de l'audience de la situation de communication. En ce qui concerne l'expertise du locuteur et son niveau de spécialisation, le sous-corpus ambiant SCA met en scène aussi bien des locuteurs académiques ou professionnels que des acteurs sociaux du troisième type qui, selon la définition de Petit (2010 : 9), partagent tous la même spécificité, celle de produire des textes en relation étroite avec le domaine du cacao et du chocolat équatorien. Ses acteurs présentent deux points communs : celui d'utiliser un environnement virtuel connecté pour produire leur discours et ils ont tous l'intention de communiquer sur le chocolat ou le cacao. Le processus de compilation du sous-corpus ambiant SCA ne permet pas d'obtenir plus d'information sur ces acteurs

du troisième type. Selon Petit, la production langagière inscrite dans un domaine spécialisé se prête à trois formes discursives : le spécialisé académique (recherche sur le domaine), le spécialisé professionnel (activité spécifique au domaine) et le spécialisé du « troisième type » (à l'exclusion de l'activité spécialisée elle-même) qui met en œuvre des acteurs liés au domaine sans être académique ou professionnel. Chacune de ces formes s'enracine dans une activité qui requiert la maîtrise de compétences particulières spécifiquement celles liées à la communication. Ils sont tous informés et possèdent suffisamment d'expertise pour que leurs discours soient crédibles dans le domaine de spécialité, même si les situations de communication prennent place hors des activités spécialisées. Le discours SCQ est en grande majorité produit par des acteurs sociaux du troisième type (Petit, 2010) considérés comme des *aficionados* suffisamment intéressés par le chocolat équatorien pour, par exemple, assister au salon du chocolat où les questionnaires ont été appliqués. Leur intention de communication est directement liée à l'expression d'un discours du sensoriel, que le stimulus soit une évocation ou une expérience directe. Le discours SCP est produit par des professionnels du domaine et l'intention de communication s'inscrit dans des écosystèmes linguistiques en relation directe avec les activités propres du domaine. Dans ce sous-corpus, les intentions de communication sont liées avec les objectifs professionnels de la vente et de la promotion.

Si la variable « niveau d'expertise » peut être considérée comme discrète, l'ampleur de l'audience, la seconde variable extralinguistique considérée, est dichotomique. Le sous-corpus SCA s'adresse à une large audience intéressée par le produit et qui se montre prête à s'informer aussi bien sur le produit que sur les domaines associés. Les discours considérés et compilés de ces situations de communication sont principalement asynchroniques et scripturaux. Les textes du sous-corpus SCA proviennent de publications réalisées sur Internet.

Les sous-corpus SCQ et SCP sont, quant à eux, compilés à partir de situations de communication mettant en scène une audience réduite au sein d'une interaction synchrone dont l'origine est à l'initiative du chercheur pour les besoins de l'étude

réalisée. Il apparaît important de prendre en compte ces variables extralinguistiques aussi bien pour l'analyse proposée que pour son interprétation (*Tableau 1* : Paramètres des situations de communication sélectionnées pour compiler le corpus d'étude).

<b>Critères</b>	<b>sous-corpus ambiant (SCA)</b>	<b>sous-corpus questionnaires (SCQ)</b>	<b>sous-corpus entretiens professionnels (SCP)</b>
<b>Canal de production</b>	Écrit	Écrit (instructions orales)	Oral (guide d'entretien écrit)
<b>Canal d'énonciation</b>	Écrit	Écrit	Oral
<b>Support d'analyse</b>	Textes originaux	Textes originaux	Textes transcrits
<b>Temporalité</b>	Pérenne	Pérenne	Éphémère
<b>Contexte social</b>	Domaine de spécialité sur Internet	Salon du chocolat Cumbayá – 2019 et Puyo en Amazonie - 2019	Différents lieux de travail - 2019
<b>Interprétant</b>	Internaute	Adultes équatoriens	Adultes équatoriens
<b>nombre</b>	Grande audience	Deux interlocuteurs	Deux interlocuteurs
<b>présence</b>	Absent lors de la production du texte	Présent lors de la production du texte	Présent lors de la production du texte
<b>Type d'interaction</b>	Indirecte et asynchrone	Directe et synchrone	Directe et synchrone
<b>Auteur</b>	Du troisième type (Petit, 2010) et professionnels du secteur	<i>Aficionados</i>	Professionnels du secteur
<b>Énonciateur</b>	Du troisième type (Petit, 2010) et professionnels du secteur	<i>Aficionados</i>	Professionnels du secteur
<b>Factualité</b>	Information factuelle	Opinions personnelles	Information professionnelle
<b>Fonction</b>	Informier	Décrire, évaluer et partager	Décrire, faire la promotion et convaincre
<b>Thème</b>	Le cacao et le chocolat équatorien	Dégustation de chocolat	L'entreprise et ses produits (cacao / chocolat)
<b>Situation géographique</b>	Publications sur Internet	Salon du chocolat Cumbayá et Puyo en Amazonie	Entreprises équatoriennes
<b>Période</b>	2005 - 2020	2019	2019

<b>Tokens</b>	1133129	5955	59617 en espagnol 6416 en kichwa
---------------	---------	------	-------------------------------------

*Tableau 1 : Paramètres des situations de communication sélectionnées pour compiler le corpus d'étude*

*Source : élaboré par nous*

### **Sous-corpus ambiant - SCA**

La compilation de ce sous-corpus ambiant, essentiellement prescriptif – au sens de Gautier et coll. (2015) et Mancebo-Humbert (2019, 59). Ces discours sont particulièrement évaluatifs et directifs dans la mesure où ils évaluent le produit, en diffusent une image dans le but de le différencier, puis incitent à l'achat. Peu importe que l'intention soit manifeste ou cachée, ce type de discours comporte toujours une forme d'incitation à l'achat. Pour que les termes et les concepts se diffusent, le discours prescriptif joue un rôle crucial, notamment auprès des consommateurs réguliers qui peuvent l'intégrer dans leur propre discours. Les discours de la filière passent par les grandes revues, les sites internet spécialisés ou les sites internet des professionnels, l'interprofession ou les syndicats, mais aussi par les grands guides et, dans une moindre mesure, par les acteurs de la filière. Il intègre des sources différentes, mais complémentaires pour offrir une représentation sociale des stratégies discursives utilisées afin de partager des informations sur le cacao et le chocolat qui sont produits en Équateur. Pour compiler ce sous-corpus, le programme Bootcat<sup>31</sup> a été utilisé de manière à sélectionner sur Internet des sources représentatives de ce discours ambiant. Pour ce faire le programme demande une liste de mots-clés – *seeds* – (*Figure 16* : Mots-clés sélectionnés pour constituer le corpus ambiant) qui servent à générer des associations de mots-clés – *tuples* – (*Figure 17*) pour une recherche systématique sur Google© (*Figure 18*) de façon à compiler l'information des « URL » retournées (*Figure 19*) sous forme de corpus archivé en format « txt » (UTF8).

<sup>31</sup> <https://bootcat.dipintra.it/?section=home>

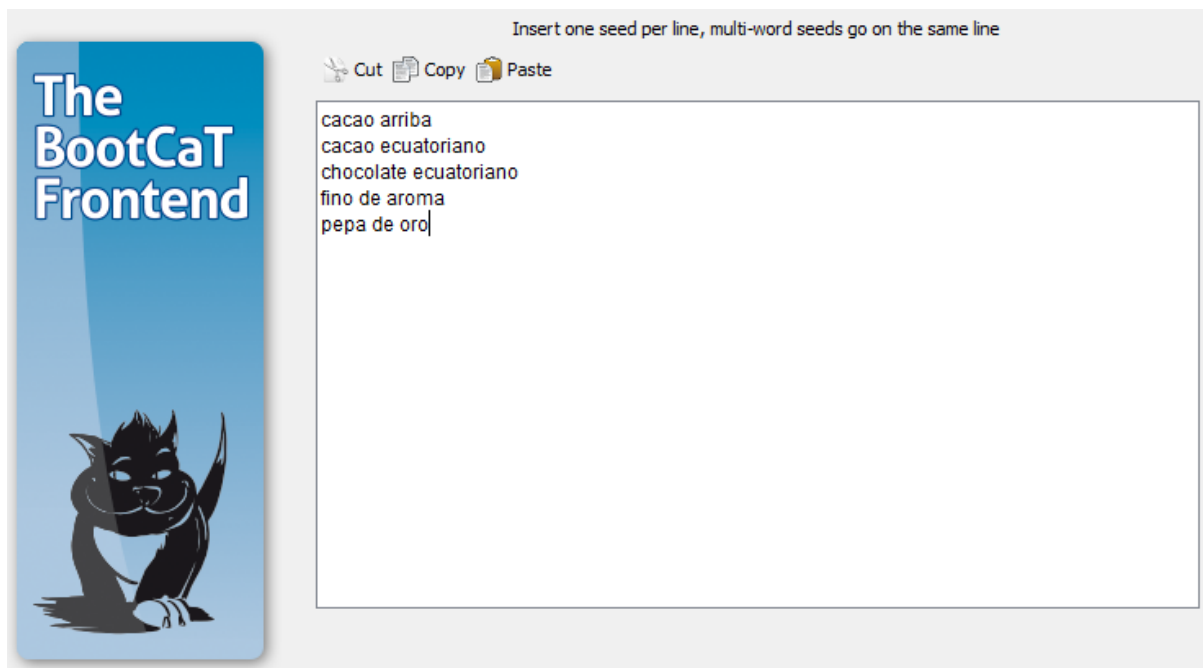


Figure 16 : Mots-clés sélectionnés pour constituer le corpus ambiant  
 Source : élaboré par nous

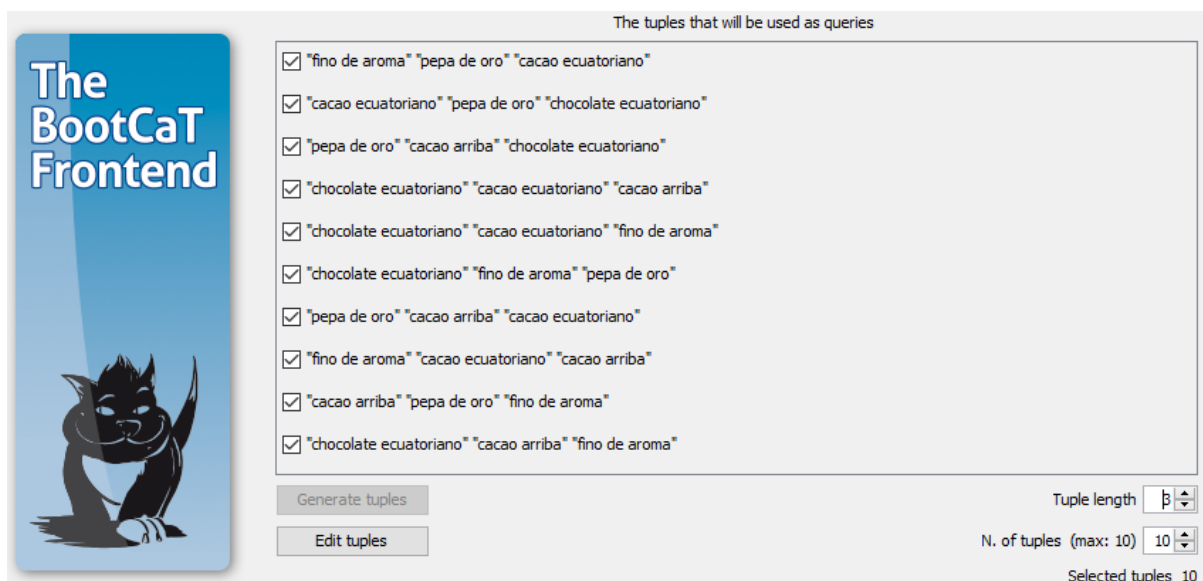


Figure 17 : Association de mots-clés sélectionnés pour faire une recherche sur Internet  
 Source : élaboré par nous

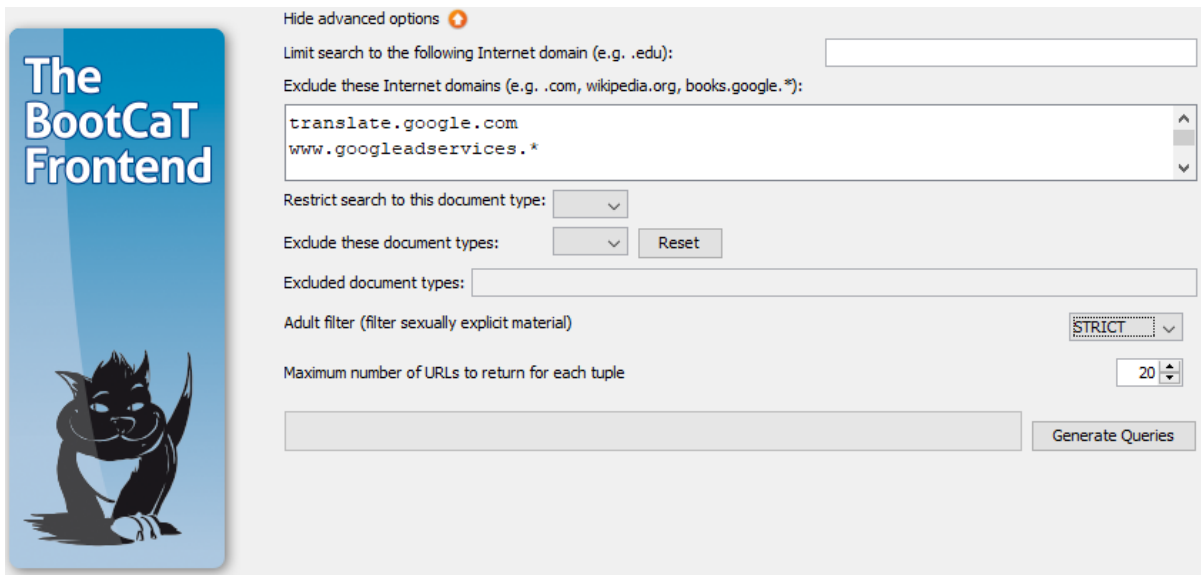


Figure 18 : Paramètre de la recherche sur Internet pour compiler le corpus ambiant  
 Source : élaboré par nous

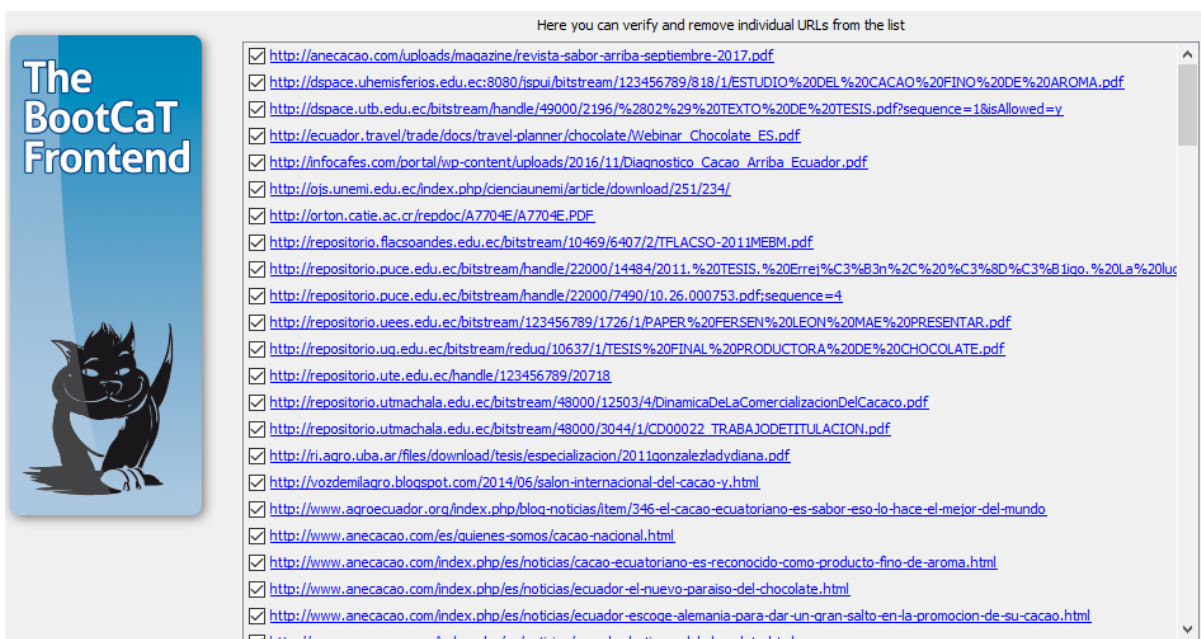


Figure 19 : Exemples d'URL sélectionnés dans Bootcat pour compiler le corpus ambiant  
 Source : élaboré par nous

L'Annexe 1 : Sources et paramètres des textes compilés pour le sous-corpus Ambient permet de visualiser les paramètres des textes compilés à partir du logiciel Bootcat ainsi que l'extension du sous-corpus dont voici un exemple :

*C-006 : El cacao fino y de aroma tiene características distintivas de aroma y sabor buscadas por los fabricantes de chocolate. Representa únicamente 5% de la producción mundial de cacao. Ecuador, por sus condiciones geográficas y su riqueza en recursos biológicos, es el productor por excelencia de Cacao Arriba fino y de aroma (63% de la producción mundial) proveniente de la variedad Nacional cuyo sabor ha sido reconocido durante siglos en el mercado internacional. Este tipo de grano es utilizado en todos los chocolates refinados. Sin embargo, los que muchos no saben que el chocolate fino se distingue por su pureza, específicamente, el sabor y fragancia que el cacao tiene.*

La raison d'être de ce sous-corpus dans cette étude, est de pouvoir compter sur une source authentique et représentative d'un discours général situé dans l'écologie linguistique équatorienne qui informe sur le domaine de spécialité du cacao et du chocolat de façon à identifier les descripteurs spécifiques au discours sensoriel utilisés en espagnol pour décrire le chocolat en contexte équatorien.

### **Sous-corpus questionnaires - SCQ**

Ce sous-corpus représentatif du discours de l'*aficionado* a été compilé à partir d'enquêtes réalisées, d'une part, auprès de consommateurs potentiels présents au salon du chocolat à Cumbayá près de Quito, au cœur des Andes équatoriennes ; et d'autre part, auprès de consommateurs potentiels de la province de Pastaza. Ces enquêtes ont été divisées en deux parties successives : en premier lieu un questionnaire écrit sans stimulus (*Annexe 2 : Questionnaire sans stimulus*), suivi d'un questionnaire écrit avec stimulus (*Annexe 3 : Questionnaire avec stimulus*). Le premier, considéré comme cognitif, se réfère aux questions visant à identifier la valeur sémantique des descripteurs



à partir d'un processus d'abstraction progressive (Dubois, 2010). Le deuxième, considéré comme empirique, se réfère aux questions qui suivent la dégustation d'un morceau de chocolat soit à 70%, soit à 100 % de cacao de façon à représenter la valeur sémantique des descripteurs à partir d'un processus de mise en discours des sensations et des émotions ressenties. Cette seconde partie s'inspire de la méthodologie mise en place dans les recherches de Mancebo-Humbert (2019). Les instructions pour remplir les questionnaires ont été données oralement.

Un total de 130 enquêtes a été appliqué dont 91 à partir d'un chocolat à 70% et 39 avec un chocolat à 100 % de cacao<sup>32</sup> dont les caractéristiques sont présentées dans le *Tableau 2* : Paramètres du sous-corpus questionnaire.

Exemple du sous-corpus compilé à partir du questionnaire sans stimulus :

*C-007 : Pienso en una taza de chocolate caliente en la tarde, una barra de chocolate semiamargo, bombones / Relajante, en mi época de niño siempre me gustaba los que son comerciales Manicho, ahora ya tenemos propios procesados y hay diversos productos*

Exemple du sous-corpus compilé à partir du questionnaire avec stimulus :

*C-008 : Estoy pensando que es muy artesanal ya que tiene sabor único, el olor impresionante y satisfactorio / Fuerte Es de mayor calidad frente a los comunes Siento el cacao purito*

<b>Critères</b>	<b>Situation de communication étudiée</b>
<b>Canal de production et d'énonciation</b>	Oral écrit selon le concept de continuum de Koch et Oesterreicher (2001)
<b>Trace temporelle</b>	Discours pérenne

<sup>32</sup> Il est important de clarifier que d'autres sessions de collecte d'information avaient été planifiées, mais l'urgence sanitaire aussi bien au niveau national qu'international, due à la propagation de la COVID-19, a limité les possibilités de réaliser ce type d'activité.

<b>Support</b>	Produit sous forme de texte	
<b>Nature</b>	questionnaire ouvert	
<b>Cadre</b>	Contexte professionnel et promotionnel du Salon du chocolat à Quito en 2019	
<b>Destinataire</b>	Adultes volontaires qui visitent le salon	
	<b>Pluralité</b>	Un individu à la fois
	<b>Présence</b>	production synchronique
	<b>Interaction</b>	Directe (lecture des questions et écriture des réponses en présence du chercheur)
<b>Auteur</b>	professionnels et <i>aficionados</i>	
<b>Fonctions</b>	Décrire et évaluer	
<b>Thème</b>	Les émotions autour du chocolat	
<b>Localisation</b>	Salon du chocolat Quito - Equateur	
<b>Période</b>	Samedi 7 et Dimanche 8 juillet 2019	
<b>Nom du Corpus</b>	Questionnaire <i>aficionado</i>	

<b>Sous-corpus</b>	<b>Corpus complet</b>	<b>Sans stimulus</b>	<b>stimulus à 70%</b>	<b>stimulus à 100%</b>
<b>Token</b>	5955	2870	2300	785
<b>lexique</b>	1153	723	624	306

Tableau 2 : Paramètres du sous-corpus questionnaire

Source : élaboré par nous

### **Sous-corpus entrevues professionnels - SCP**

Ce sous-corpus a été compilé à partir d'entretiens semi-dirigés (*Annexe 4* : Guide d'entretien pour les entrevues en espagnol et *Annexe 5* : Guide d'entretien pour les entrevues en kichwa) de professionnels impliqués dans la chaîne de production de la fabrication de chocolat majoritairement situés dans la région amazonienne de l'Équateur, dans les provinces de Pastaza et de Napo selon la distribution suivante :

#### **Province Pastaza :**

- Canton Santa Clara : Communauté Handuayacu, San Jorge et San Rafael

- Canton Arajuno : Communauté Boano, Nushiño Ishpingo, San Juan de Shiculi, Shiguarcocha et Helena Andi
- Canton Pastaza : Communauté Rey de Camino, Porvenir et Esperanza.

## **Province Napo**

- Cantón Tena: Comunidad Shandia, Serena, Centro Talag, et San pablo

Dans les communautés citées, soixante-deux entretiens ont été réalisés auprès des cultivateurs de cacao grâce à l'aide des techniciens du MAGAP, institution provinciale de l'équivalent du ministère de l'Agriculture en Équateur. Quarante-trois de ces entretiens se sont effectués en espagnol et dix-neuf en kichwa. Ces acteurs sociaux vivent dans des communautés en Amazonie, la grande majorité ont le Kichwa comme langue native. Comme tout élément naturel, la fève de cacao est traditionnellement présente dans leur écosystème linguistique contrairement à la barre de chocolat connue des pays consommateurs. Ils ont pour habitude de consommer le cacao sous forme de boisson. Cependant, pour répondre à la demande externe, ils commencent à fabriquer du chocolat en barre le plus souvent très concentré, entre 70 et 100 % de cacao.

Ce sous-corpus a été complété par les entretiens dirigés aux professionnels qui achètent, transforment le cacao pour, fabriquer et vendre le chocolat. Vingt-sept entretiens ont été enregistrés en espagnol avec les responsables d'entreprises productrices de chocolat comme Saquifracia qui est une petite entreprise familiale située à Puyo, où ils cultivent, traitent et fabriquent le chocolat. Comme le cacao qu'ils produisent ne suffit pas pour leur production de chocolat, ils sont obligés d'acheter du cacao à d'autres exploitations. Ils élaborent plusieurs produits dérivés à base de cacao comme la crème de cacao ou le vin *espumante* (mousseux) fabriqué à partir de la fermentation lactique de la pulpe du cacao.

Autre exemple d'entreprise sélectionnée pour les entretiens, Kallari qui en kichwa signifie : initier, ou encore le début de quelque chose. Plus qu'une entreprise, c'est une

association, une coopérative, composée de 850 producteurs de cacao, dont la majorité est composée de familles kichwa de 21 communautés du canton Tena de la province de Napo de l'Amazonie équatorienne. Cette association produit, transforme et commercialise des produits organiques qui respectent le commerce équitable et durable permettant ainsi d'améliorer les conditions de vie de ceux qui font partie de l'association tout en conservant la biodiversité naturelle et le patrimoine culturel. Les autres professionnels qui ont collaboré aux entretiens sont identifiés en *Annexe 6* : Paramètres des textes compilés pour le sous-corpus professionnel producteur de chocolat.

Un total de quarante-et-un enregistrements ont été compilés et transcrits en espagnol (*Annexe 7* : Paramètres des textes compilés en espagnol pour le sous-corpus professionnel producteurs de cacao) et un total de vingt enregistrements en kichwa (*Annexe 8* : Paramètres des textes compilés en kichwa pour le sous-corpus professionnel producteur de cacao). Les paramètres décrits dans le *Tableau 3* : Paramètres du sous-corpus professionnel permettent de lier le sous-corpus à son écosystème linguistique.

<b>Paramètres</b>		<b>Situation de communication étudiée</b>
<b>Canal de production et d'énonciation</b>		Oral
<b>Trace temporelle</b>		Discours éphémère
<b>Support</b>		Discours oral enregistré et transcrit
<b>Nature</b>		Entrevue semi-dirigée
<b>Cadre</b>		Contexte professionnel et promotionnel du Salon du chocolat à Quito, Coopérative Kallari à Tena en 2019, et fermes de production du cacao en Amazonie
<b>Destinataire</b>		Adultes volontaires impliqués dans la chaîne de production du chocolat
	<b>Pluralité</b>	Un individu à la fois
	<b>Présence</b>	production synchronique
	<b>Interaction</b>	Directe (lecture des questions du guide d'entrevue et réponses directes)
<b>Auteur</b>		professionnels
<b>Fonctions</b>		Décrire, faire la promotion, et convaincre
<b>Thème</b>		Les émotions autour du chocolat
<b>Localisation</b>		Sur le lieu de travail en Équateur

<b>Période</b>	Samedi 7 et Dimanche 8 juillet 2019 / 22 janvier 2019
<b>Nom du Corpus</b>	sous-corpus professionnel (SCP)

<b>sous-corpus</b>	<b>Producteur de chocolat</b>	<b>Producteur de cacao - espagnol</b>	<b>Producteur de cacao -kichwa</b>
<b>Token</b>	30934	28683	6416
<b>Vocabulaire</b>	3217	2172	1491
<b>Temps audio</b>	3h12'05"	02h33'30"	01h11'30"

Tableau 3 : Paramètres du sous-corpus professionnel

Source : élaboré par nous

Étant compilé à partir d'entretiens, le sous-corpus professionnel est construit avec les transcriptions des enregistrements réalisés. Pour être fidèle aux caractéristiques de l'oralité, la transcription doit éviter d'être interprétative et ainsi permettre un accès objectif à la réalité de la situation de communication enregistrée. De façon à respecter l'écosystème linguistique de la production orale, les transcriptions ont été réalisées par des personnes appartenant à cet écosystème. Ce choix a pour conséquence la présence d'erreur de langage qui illustrent les compétences langagières de ces acteurs sociaux qui: ils écrivent *osea* en un seul mot au lieu de *o sea*, *haber* au lieu de *a ver*.

Il reste inévitable que pendant l'exercice de la transcription de nombreux éléments communicatifs de l'oralité se perdent. Comme le souligne Linell, il existe un certain désavantage à transformer un processus dynamique oral en un processus statique textuel (2005). Cependant, la transcription reste le seul exercice possible pour matérialiser l'oralité dans un format exploitable par les outils de la textométrie. Le défi principal de la transcription est d'offrir une description suffisamment fidèle de la réalité discursive orale tout en évitant un niveau de complexité trop élevé qui rend difficiles la lecture et l'accès à l'information (Recalde et Rozas, 2009). Afin de relever ce défi tout au long du processus de transcription, il est recommandable d'établir des normes de transcription pour limiter les variations synchroniques, entre différents transpositeurs, et asynchroniques, d'un même transpositeur sur plusieurs enregistrements. Les entretiens ont été enregistrés et transcrits en s'inspirant des recommandations du protocole de

transcription orthographique COSER - Corpus Oral y Sonoro del Español Rural (Bejarano et coll., 2018) selon le protocole décrit dans le *Tableau 4* : Paramètres de transcription :

Données générales	
entrevistado :	La personne qui répond aux questions
entrevistador :	La personne qui dirige l'entretien
¿?	Questions
¡!	Exclamations
<eh>	Onomatopées
<ah>	
<oh>	
<m>	
<...>	Partie de l'enregistrement inaudible
<aa>	forme d'acquiescement
<jajaja>	Rires
<jejeje>	
Orthographe	
Le transcripteur, qui appartient à l'écosystème linguistique du participant, écrit ce qu'il entend selon les compétences linguistique qu'il possède. Il écrit les mots tronqués et les formes répétées.	
Mis à part les points d'interrogation et d'exclamation, aucun autre signe de ponctuation n'est utilisé. Le transcripteur va à la ligne à chaque fin de contribution.	

*Tableau 4 : Paramètres de transcription*

*Source : élaboré par nous*

Afin d'illustrer l'utilisation des paramètres de transcription, nous proposons deux exemples :

*C-009 : **entrevistador:** se dedicó al cacao bien y ¿hace cuántos <eh> hace cuánto tiempo que se dedica a la producción de cacao?*

***entrevistado:** ya son unos treinta años o más*

***entrevistador:** treinta años o más bien y ¿puede definir las características del chocolate elaborado con su cacao?*

***entrevistado:** <m> no sé cómo explicarle*

*no he tenido mucho conocimiento que no más hacen*

*se que hacen chocolate*

*pero de ahí no se los dulces así  
caramelos que está incluido ahí con el chocolate mismo entonces  
nada más esas cosas que he conocido*

**entrevistador:** *no ha hecho entonces chocolate con su cacao*

**entrevistado:** *no no no hecho*

**entrevistador:** *bien le agradezco muchísimo sí*

C-010 : **entrevistador:** *que es características <m> ¿cómo era?*

*¿cómo era el chocolate?*

**entrevistado:** *primerito era así como agrio*

**entrevistador:** *<ah> ya agrio*

**entrevistado:** *ya si*

*de ahí ya pusimos az azúcar leche*

*de ahí se quedó más o menos ya*

**entrevistador:** *ya y que*

**entrevistado:** *estaba rico el chocolate*

**entrevistador:** *estaba rico ya*

*y ¿qué sabor tiene este chocolate?*

*¿qué sabor tenía este chocolate?*

**entrevistado:** *sabor a cacao <jajaja>*

**entrevistador:** *sabor a caco*

*y ¿a qué le le hizo pensar?*

**entrevistado:** *<m> <jajaja> <m> <...> <jajaja>*

**entrevistador:** *¿en qué pensó cuándo probó el chocolate?*

**entrevistado:** *en estaba bueno*

**entrevistador:** *que estaba bueno*

**entrevistado:** *y pensé en hacer más cacao*

*y vender chocolate yo mismo preparando <jajaja>*

*entrevistador: bien*

*entrevistado: eso pensé*

*entrevistador: y ¿qué sensación tie tuvo cuando probó el chocolate?*

*entrevistado: <...> <jajaja> y qué sensación*

*entrevistador: ¿qué sintió?*

*entrevistado: bien*

*entrevistador: ¿qué sensación?*

*entrevistado: feliz bien <jajaja>*

Les questions de l'entretien ont été incluses dans la transcription et compilées dans le corpus, car il a été constaté de nombreux cas de reformulation des questions pour adapter l'entretien aux différentes réponses des participants. Conscient que ce choix a de fortes répercussions sur l'analyse textométrique, notamment sur la fréquence d'utilisation des mots employés dans les questions et les cooccurrences, il est important de tenir compte de cette particularité lors des interprétation des résultats.

En résumé, ce chapitre présente le concept de linguistique sensorielle dans la communication des vertus du cacao et des produits à base de chocolat. Il explore également le concept de discours sensoriel et sa signification en relation avec l'émotion, la perception et l'interprétation. Il aborde le processus de compilation d'un corpus pour une étude linguistique sur le discours sur le chocolat et explique les critères de sélection des textes pour le corpus. Il se penche également sur les variations des critères de sélection des sous-corpus et sur les méthodes utilisées pour la collecte des données. Le chapitre conclut en discutant des défis et des recommandations pour la transcription des entretiens oraux en vue de leur analyse.

Le thème de la linguistique sensorielle est considéré pour sa pertinence avec les intentions de communication des acteurs de l'industrie du cacao et du chocolat. Ce positionnement théorique de la recherche s'inscrit dans la lignée des travaux menés par



l'équipe du professeur Gautier au sein du laboratoire TIL (Texte, Image, Langage) à Dijon, France. En effet, le rôle de la linguistique sensorielle dans la communication des vertus d'un produit ou d'un service, en particulier dans le contexte de l'industrie du cacao et du chocolat, est aussi bien lié aux cinq sens impliqués dans la satisfaction du consommateur qu'aux sensations multimodales stockées dans la mémoire à long terme. Le concept de linguistique sensorielle combine les neurosciences cognitives et l'analyse linguistique tout en intégrant l'utilisation de l'analyse de corpus et du traitement automatisé du langage. A travers le prisme d'une linguistique outillée, le chercheur peut analyser la construction du discours basée sur la réalité culturelle et l'intention de la communication pour exprimer à la fois les expériences sensorielles et les perceptions cognitives. Le corpus d'étude est représentatif d'un discours défini comme une manifestation dynamique et cognitive d'un message situé, immergé dans un écosystème socioculturel dont la signification est en étroite relation avec l'émotion, la perception et l'interprétation.

La linguistique outillée repose sur la compilation d'un corpus qui est une collection de textes sélectionnés et organisés sur la base de critères linguistiques et extralinguistiques pour représenter un usage spécifique de la langue dans une situation de communication particulière. Après avoir exploré différentes définitions d'un corpus, l'auteur conclut qu'un corpus est un groupe structuré de textes complets, documentés et potentiellement enrichis d'annotations, choisis de manière théorique et représentant les phénomènes linguistiques associés à une praxis sociale spécifique culturellement ancrée dans son écosystème de production langagière. A la lumière de cette définition, trois sous-corpus sont compilés selon : l'expertise du locuteur et l'étendue de l'audience dans la situation de communication. Le sous-corpus SCA comprend des textes produits par des locuteurs académiques ou professionnels ainsi que par des acteurs sociaux liés au domaine du cacao et du chocolat équatorien. Le sous-corpus SCQ se compose principalement de textes produits par des acteurs sociaux connus sous le nom d'"aficionados" qui s'intéressent au chocolat équatorien. Le sous-corpus SCP comprend des textes produits par des professionnels dans le domaine de la production du chocolat. Dans sa dernière

partie, le chapitre explique le processus de compilation et les détails sur le contexte et la méthodologie utilisés pour la collecte des données incluant les normes de transcription considérées.

Dans l'ensemble, le lecteur possède une vue d'ensemble de la linguistique sensorielle et de son application dans l'industrie du cacao et du chocolat, ainsi que de la méthodologie utilisée de l'analyse socio-discursive appliquée. En conclusion, ce chapitre nous a permis de préciser le modus operandi utilisé pour la compilation des sous-corpus, ainsi que leurs caractéristiques et leur représentativité selon les différents écosystèmes linguistiques sélectionnés auxquels ils appartiennent.

# Deuxième Partie :

## Caractérisation analytique du corpus et modélisation

---

Les différentes étapes de la première partie permettent au chercheur de disposer d'un contexte d'étude spécifique défini comme l'écosystème linguistique dans lequel les acteurs impliqués dans le domaine spécialisé du cacao et du chocolat interagissent. Ces interactions laissent des empreintes linguistiques qui, une fois compilées en corpus et sous-corpus, offrent au chercheur un accès aux stratégies de communication utilisées par les différents acteurs selon leurs intentions de communication. L'analyse et l'interprétation de ces empreintes de communication sont la raison d'être de la textométrie. En effet, en accord avec Gledhill, Patin et Zimina, « [l]a textométrie regroupe l'ensemble des méthodes quantitatives permettant d'opérer des réorganisations formelles de la séquence textuelle et des analyses statistiques portant sur l'ensemble des unités textuelles d'un corpus » (2017 : 5). Aussi connue comme « linguistique outillée », la particularité de l'analyse textométrique est d'utiliser des outils informatiques afin d'offrir une grille de lecture d'un ensemble de textes pour en interpréter les caractéristiques prototypiques. En effet, la modélisation du corpus étudié s'obtient en premier lieu par l'interprétation qualitative de données quantitatives en relation directe avec l'écosystème linguistique dans lequel les discours se produisent, approche abductive conduite par le corpus d'étude (bottom-up). Cette interprétation est ensuite vérifiée et validée à l'aide de raisonnements inductifs ou déductifs (Top-down) de façon à proposer une modélisation représentative du moule discursif de la situation de communication étudiée. Les analyses quantitatives éclairent le chercheur sur les caractéristiques des textes considérés à partir des données issues des corpus suivant une dimension heuristique de va-et-vient entre empirie et théorie appelée approche « bottom-up / top-down » (Brown et Yule, 1983 : 234–235). Dans un cadre de sémiotique peircienne, ce va-et-vient crée un rapport étroit entre l'interprétant, qui n'est autre que le chercheur, l'objet, ou encore le corpus qui représente le discours dans son écosystème linguistique, et le *representamen* que construit l'interprétant à partir de son analyse : le

prototype discursif aussi appelé moule discursif dans la tradition germanique (Fix, 1990, 1999 ; Gautier, 2009). Le concept de moule discursif peut être illustré par une jauge textuelle idéale et abstraite facilitant l'étalonnage des productions langagières de chaque sous-corpus. Commun aux prototypes que peuvent avoir les interprétants le moule discursif est un des constituants de l'écosystème sociolinguistique.

Le troisième chapitre a la prétention d'offrir au lecteur un aperçu du moule discursif du corpus considéré et de ses sous-corpus et ainsi évaluer l'empreinte du discours sensoriel dans les situations communicatives situées dans leur écosystème linguistique. La valorisation des résultats obtenus est exposée dans le cinquième et dernier chapitre de ce travail.

## Chapitre 3 : Analyse textométrique et interprétation croisée des corpus considérés

---

Après avoir compilé les sous-corpus selon les processus décrits dans le chapitre précédent, la seconde étape de la méthodologie appliquée met en œuvre les outils textométriques suivants : TXM<sup>33</sup>, TreeTagger<sup>34</sup> et Orange<sup>35</sup> qui sont en accès libre sur Internet. La plateforme TXM permet l'analyse de corpus structurés et annotés (Heiden, 2010 ; Heiden et coll., 2010 ; Pincemin et coll., 2010) en offrant la possibilité au chercheur de construire ses sous-corpus directement dans la plateforme, de les annoter en intégrant directement les fonctionnalités de TreeTagger et ainsi de compter sur un corpus lemmatisé et étiqueté selon les parties du discours. La lemmatisation permet de s'affranchir des différentes flexions des tokens, par exemple, le lemme *estar* représente toutes les flexions utilisées dans la conjugaison de ce verbe, pour un adjectif, le singulier masculin représente les flexions du genre et du nombre. Grâce au moteur de recherche intégré – Corpus Query Processor (CQP) – dans les outils présentés, le chercheur a accès aux occurrences de mots, d'expressions, de parties du discours, de chaînes de plusieurs mots (n-gram). Il peut aussi obtenir les concordances et les cooccurrences de ces requêtes, tout en ayant accès à une analyse croisée entre sous-corpus pour déterminer les statistiques spécifiques des occurrences. Les concordances sont un outil analytique de la linguistique de corpus qui permet d'examiner l'utilisation d'un mot dans différents contextes au sein d'un corpus ou d'un groupe de textes. Elles fournissent des informations sur la fréquence, la distribution et l'utilisation sémantique d'un mot dans le texte. Comme le décrivent McEnery et Wilson, les concordances sont des extraits du contexte textuel dans lequel le mot-clé apparaît (2001 : 21). Les mots-clés peuvent être n'importe quel terme individuel ou combinaison de termes que le chercheur souhaite

---

<sup>33</sup> <https://txm.gitpages.huma-num.fr/textometrie/?lang=en&gt>

<sup>34</sup> <https://cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

<sup>35</sup> <https://orangedatamining.com>

analyser. Par exemple, pour étudier l'utilisation du verbe *gustar* dans un ensemble particulier de textes, on peut créer une liste de concordances pour ce mot-clé spécifique. Ces listes comprennent généralement des morceaux de texte suffisamment long pour offrir un contexte dans lequel le mot-clé apparaît, ainsi que certains éléments grammaticaux environnants afin de fournir une interprétation sémantique plus complète. Il peut s'agir d'informations sur le nombre, le genre et la personne du sujet et de l'objet autour du verbe. En résumé, d'un point de vue linguistique à travers la linguistique de corpus, les concordances sont essentiellement des mots-clés obtenus par une recherche informatique contextualisée dans des corpus textuels dont l'objectif est de mieux comprendre comment certains mots ou certaines structures syntaxiques s'utilisent. Dans la linguistique de corpus, la cooccurrence fait référence à l'apparition simultanée de deux mots ou plus dans un contexte spécifique. L'analyse des cooccurrences permet d'extraire des modèles significatifs en termes de sémantique et de syntaxe qui facilitent le processus d'interprétation du corpus. Selon Sinclair, la cooccurrence est une propriété fondamentale du langage naturel, où les mots acquièrent leur signification à partir de modèles contextuels (1991 : 123). Cela signifie que le sens d'un mot est déterminé par les mots qui apparaissent dans son environnement. Par exemple, si nous analysons la fréquence à laquelle les mots *pepa* et *oro* apparaissent ensemble dans un corpus relatif au chocolat en Equateur, nous pouvons en déduire que ces deux mots sont étroitement liés sur le plan sémantique au point de former une expression figée – *pepa de oro* – pour décrire la fève de cacao. En outre, cette relation permet de formuler des hypothèses sur d'autres aspects comme le figement ou non d'une expression, ou encore s'il existe des différences régionales ou culturelles. L'analyse des cooccurrences peut être utile pour comprendre comment les locuteurs utilisent différents termes et expressions dans un écosystème discursif particulier. Elle peut également être utilisée pour mieux comprendre comment certaines associations lexicales sont courantes dans différents genres textuels. En résumé, d'un point de vue linguistique, par le biais de l'analyse de corpus, la cooccurrence se réfère aux relations entre les éléments lexicaux compte tenu de leur proximité contextuelle, dont l'interprétation nous permet d'approfondir non seulement les aspects sémantiques, mais

aussi syntaxiques présents dans le discours.. La classification ascendante hiérarchique (CAH) est une méthode de classification automatique qui mesure la dissimilarité entre les individus d'un groupe de façon à connaître l'organisation des données. Dans le cas de la textométrie, les CAH permettent de connaître la distance entre les différents éléments d'un sous-corpus et d'en déduire leur affinité ou leur divergence. Elles classifient les individus dans n-classes restituées dans une représentation graphique sous forme de dendrogramme qui illustre le grade de dissimilarité, ou d'affinité, selon la distance qui sépare les individus.

Dans cette étude, l'analyse des variable précédemment décrites offrent des données quantitatives qui permettent au chercheur et un accès direct à une source primaire d'information. Une fois resituée dans l'écosystème linguistique de production des textes, cette source lui permet de proposer une interprétation qualitative du vocabulaire utilisé dans les différentes situations de productions langagières considérées en relation au lexique de l'écosystème sociolinguistique étudié. En effet, même si les termes « vocabulaire » et « lexique » sont souvent utilisés de manière interchangeable pour désigner l'ensemble des termes qui composent une langue, il existe certaines différences entre ces concepts. Dans cette étude, en accord avec l'approche de Aitchison (2012), le lexique désigne spécifiquement l'ensemble des termes d'une langue appartenant à un écosystème sociolinguistique spécifique : le domaine contemporain du chocolat et du cacao en Equateur. Le lexique peut également inclure des expressions idiomatiques et des phrases qui font partie du langage courant. D'autre part, le vocabulaire fait référence au nombre limité de termes qu'un individu ou un groupe social connaît du lexique de référence. En d'autres termes, alors que le lexique comprend tout ce qui existe dans un écosystème sociolinguistique - à la fois connu et inconnu - le vocabulaire est simplement ce que quelqu'un, ou un groupe d'individus, connaît ou utilise activement pour communiquer. Même si tous les individus ont potentiellement accès au même lexique, chaque individu, ou groupe d'individus, possède son propre vocabulaire personnalisé en fonction de son niveau d'expertise, de ses intérêts spécifiques ou de son contexte culturel et social.

Le programme Orange, développé par l'université de Ljubljana, offre de nombreux outils d'exploration de données tout en restituant une organisation et une visualisation intelligente. Il permet une exploration approfondie des distributions statistiques aussi bien à partir de données numériques que textuelles (corpus de texte) selon le flux d'analyse automatique que détermine le chercheur qui, une fois établi, peut être appliqué à différents corpus comme l'illustre le flux appliqué au différents sous-corpus de cette analyse (Figure 20 : Flux d'analyse des sous-corpus (Orange)).

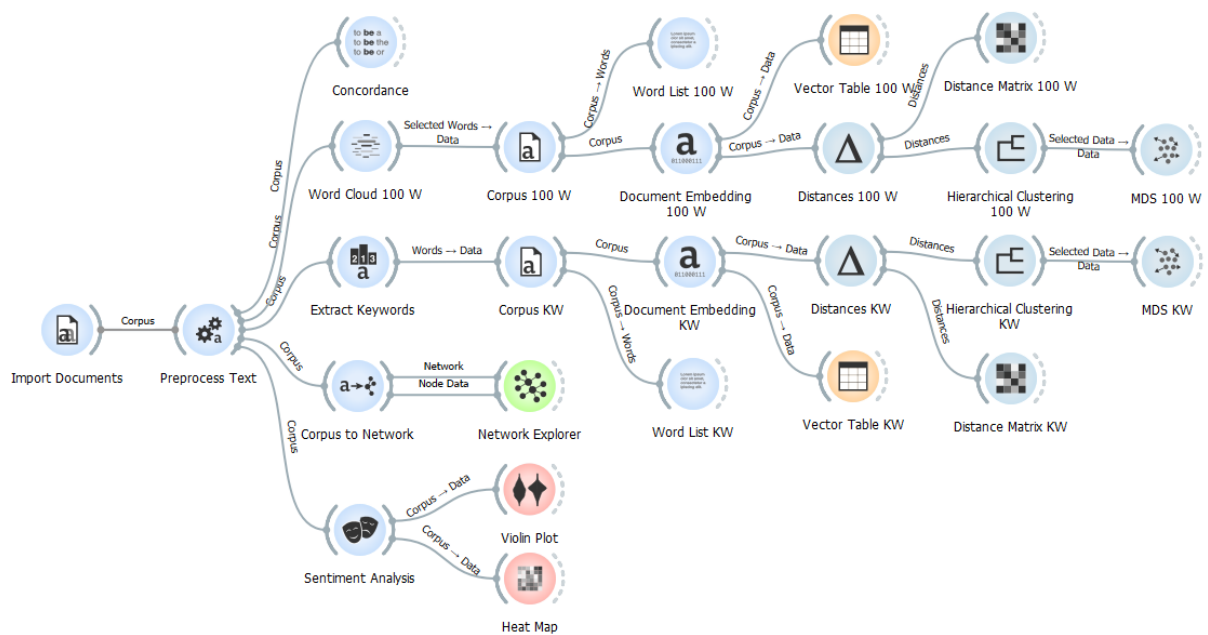


Figure 20 : Flux d'analyse des sous-corpus (Orange)  
 Source : élaboré par nous

En effet, une fois le flux de travail établi, le programme permet de charger le corpus choisi pour l'analyse. Il applique automatiquement chaque étape du processus permettant d'obtenir les mêmes types de résultats pour chacun des corpus étudiés. Dans le flux considéré, les sorties disponibles sont : *concordance*, *world cloud*, *extract*



*keywords, word list, vector table, distance matrix, hierarchical clustering, MDS, topic modelling, network explorer, violin plot et heat map*<sup>36</sup>.

L'objectif de ce troisième chapitre est de suivre une méthodologie analytique d'exploration de données qui à partir d'un corpus représentatif pour offrir une description des moules discursifs des sous-corpus considérés tout en respectant l'environnement écosystémique décrit en première partie. Une fois le moule discursif du sous-corpus ambiant présenté, une approche, selon les différentes intentions de communication des professionnels et des *aficionados*, sera proposée à partir des mêmes outils textométriques mettant en évidence leurs stratégies discursives pour exprimer leurs passions, leurs émotions ou leurs absences d'émotions quand il parle du cacao ou du chocolat.

### 3.1 Moule discursif représentatif du sous-corpus discours *ambient*

Une fois compilé, le sous-corpus ambiant (SCA) est analysé grâce aux outils textométriques disponibles. Le *Tableau 5 : Paramètres du sous-corpus ambiant* présente les informations générales mettant en évidence l'extension (1133129 tokens), garante de la représentativité des stratégies discursives utilisées afin de partager des informations sur le cacao et le chocolat produits en Équateur. L'étendue du vocabulaire (47782 tokens différents) du sous-corpus ambiant démontre la grande variété de termes utilisés pour la description de ce domaine, offrant une idée représentative de l'étendu du lexique, ce qui se reflète dans la densité lexicale. En effet, sur 1000 tokens aléatoirement sélectionnés dans le sous-corpus, il existe 513 tokens différents ; non seulement le lexique est étendu, mais il est aussi varié dans la production langagière.

---

<sup>36</sup> Chaque processus est décrit dans ce même chapitre lors de la présentation des résultats obtenus.

sous-corpus SCA	Token	Vocabulaire	Densité lexicale sur 1000 tokens aléatoires
Nombre total	1133129	47782	0.513

Tableau 5 : Paramètres du sous-corpus ambiant

Source : élaboré par nous

De nombreux paramètres ont été définis pour calculer la richesse lexicale d'un texte. Torruella Casaña et Capsada Blanch (2017) présentent une étude de révision des différentes méthodes. Pour mesurer la richesse lexicale d'un corpus et la comparer à d'autres corpus, ces auteurs ont étudié le comportement d'une quinzaine d'indices de mesure de la richesse lexical. Considérant les critères de stabilité, de sensibilité et de cohérence (2017 : 389) ainsi que les différences de taille des sous-corpus de l'étude, le calcul du paramètre D à partir d'un échantillon aléatoire de 1000 tokens a été choisi pour pouvoir comparer la richesse lexicale des sous-corpus étudiés en limitant l'impact de la taille du sous-corpus sur le résultat. Le calcul de ce paramètre se réalise à partir d'un échantillon de mille tokens aléatoires sélectionnés dans le corpus et le nombre de mots différents que présente cet échantillon. La fonction *shuffle words* de l'outil *text converter* disponible sur Internet<sup>37</sup> permet de mélanger de manière aléatoire les mots d'un corpus pour en sélectionner une quantité significative et obtenir le rapport entre le nombre de mots de l'échantillon, arbitrairement établi à mille, et le nombre de mots différents que cet échantillon présente (échantillon consultable en *Annexe 9* : Échantillon de 1000 mots aléatoires du sous-corpus *Ambiant*). Cette méthode sera appliquée aux autres sous-corpus pour ainsi obtenir une représentation de la richesse lexicale comparable.

Dans le *Tableau 6* : 100 premiers lemmes des noms communs, adjectifs, verbes et adverbes du sous-corpus ambiant, il est possible d'observer les cent noms communs,

---

<sup>37</sup> Cet outil informatique librement disponible sur internet (<https://textconverter.com/shuffle-words>) permet la réalisation de manipulations simples d'un texte ou de son format. Parmi les fonctions proposées, il permet l'extraction de parties spécifiques comme les URLs, la possibilité d'extraire des échantillons aléatoires ou encore la suppression de la ponctuation.

adjectifs, verbes, et adverbes les plus fréquents dans le sous-corpus ambiant avec leur fréquence respective. L'algorithme utilisé qui permet l'étiquetage des tokens selon les parties du discours est directement intégré dans le programme Orange, il se situe au niveau de la fonction *Preprocess Text*. Cet algorithme, nommé Average Perceptron Tagger, a été développé par Matthew Honnibal qui affirme que cet outil d'étiquetage de texte a une précision d'environ 97 % (2013). Il est donc possible que 3 % des balises attribuées aux tokens soient erronées et que des termes comme *ecuatoriano*, *poder* soient balisés comme substantifs alors qu'ils peuvent être respectivement un adjectif et un verbe. Cependant, la marge d'erreur reste acceptable d'autant plus qu'elle est la même sur tous les sous-corpus ce qui limite son impact lors de comparaisons entre sous-corpus.

[espos="NC"] eslemma	F	[espos="ADJ"] eslemma	F	[espos="V.*"] eslemma	F	[espos="ADV"] eslemma	F
cacao	13014	grande	2141	ser	12461	más	2812
chocolate	3928	bueno	1665	haber	4081	también	1218
producto	3609	fino	1413	tener	3208	ya	891
producción	3001	internacional	1270	estar	2460	muy	775
grano	2813	ecuatoriano	1264	poder	2331	donde	770
país	2640	pequeño	1111	deber	2017	como	641
mercado	2294	mundial	1028	hacer	1459	así	611
año	2259	nacional	1023	realizar	1241	dentro	599
calidad	2151	importante	877	producir	939	tanto	571
exportación	1875	principal	835	obtener	935	además	550
productor	1803	alto	794	permitir	899	bien	422
empresa	1543	mismo	788	dar	876	menos	320
aroma	1522	orgánico	788	existir	776	siempre	291
sector	1412	nuevo	779	conocer	771	antes	271
sabor	1380	productivo	686	encontrar	771	actualmente	244
proceso	1372	productor	631	utilizar	771	luego	236
nivel	1256	nuestro	583	decir	715	después	230
precio	1221	económico	535	elaborar	648	principalmente	217
valor	1118	diferente	499	exportar	648	poco	215
cultivo	1010	total	495	ir	581	directamente	196
tipo	931	necesario	463	considerar	544	solamente	196
cadena	887	comercial	456	generar	523	ahora	192
mundo	858	siguiente	427	lograr	519	hoy	191

variedad	829	último	405	presentar	511	solo	191
forma	777	primo	400	seguir	510	tan	188
origen	775	social	377	ver	503	cómo	185
desarrollo	774	solo	370	establecer	499	entonces	185
fermentación	732	bajo	367	secar	492	mucho	185
industria	682	agrícola	361	mejorar	490	cuanto	173
agricultor	681	local	335	desarrollar	472	aproximadamente	170
muestra	665	natural	332	representar	472	cerca	146
actividad	648	interno	326	llegar	469	aquí	144
comercialización	637	seco	318	mantener	464	aún	134
investigación	634	público	317	trabajar	449	especialmente	134
comercio	631	técnico	314	contar	448	finalmente	132
análisis	621	práctico	303	sembrar	431	casi	123
trabajo	603	final	295	tomar	404	incluso	106
organización	601	general	295	buscar	398	obstante	99
manera	600	presente	284	comprar	394	frente	89
marca	595	especial	280	pasar	362	ahí	84
exportador	588	propio	273	cumplir	359	realmente	76
condición	584	único	269	reconocer	354	posteriormente	75
información	582	exportador	268	incluir	345	asimismo	74
zona	580	anual	265	determinar	331	No	74
tonelada	577	actual	242	llevar	331	únicamente	74
asociación	576	competitivo	242	indicar	326	adicionalmente	71
planta	570	directo	238	afectar	322	pesar	71
materia	559	externo	238	aplicar	320	adelante	70
objetivo	544	potencial	235	alcanzar	318	generalmente	70
provincia	536	posible	233	crear	308	primero	70
sistema	534	estratégico	224	aumentar	307	nunca	68
fuerza	532	sostenible	221	mostrar	299	todavía	67
demanda	524	igual	218	incrementar	292	respectivamente	65
proyecto	524	específico	211	requerir	287	básicamente	64
costo	509	libre	204	contener	283	bastante	63
consumo	494	junto	195	cultivar	281	altamente	59
servicio	491	justo	194	emplear	276	allí	57
consumidor	476	privado	194	recibir	274	atrás	55
embargo	472	floral	190	observar	266	anteriormente	54
manteca	472	tradicional	187	asegurar	264	totalmente	54
certificación	470	fuerte	184	crecer	263	específicamente	51
factor	463	geográfico	184	ubicar	261	exclusivamente	49
elaboración	461	métrico	177	querer	260	prácticamente	48
característica	457	superior	177	beneficiar	257	acerca	47

estrategia	456	industrial	176	identificar	257	demasiado	46
fin	446	europeo	171	vender	257	pese	46
plantación	435	adecuado	170	depender	253	ligeramente	45
acuerdo	430	químico	169	destacar	250	allá	44
cantidad	429	distinto	165	empezar	246	debajo	44
persona	429	arancelario	164	ofrecer	245	internacionalmente	43
resultado	428	largo	164	fermentar	239	apenas	39
relación	427	genético	161	venir	239	fuera	39
estudio	426	básico	160	usar	234	recientemente	39
cosecha	425	ambiental	158	partir	231	acá	38
tiempo	425	asociativo	153	saber	229	sí	38
barra	423	mediano	152	poner	228	simplemente	38
día	420	anterior	149	plagar	225	encima	36
norma	420	alimentario	146	pagar	223	debido	35
crecimiento	418	derivado	145	evitar	221	tradicionalmente	35
suelo	417	humano	144	reducir	221	conjuntamente	34
venta	414	global	140	convertir	220	inicialmente	34
participación	410	claro	136	poseer	220	particularmente	33
uso	405	empresarial	136	promover	219	mayormente	32
ingreso	404	relativo	136	garantizar	213	parcialmente	32
dato	401	climático	135	comercializar	212	previamente	32
lugar	392	excelente	133	corresponder	212	detrás	31
productividad	392	extranjero	132	recuperar	212	inclusivamente	31
región	392	ambiente	130	constituir	210	normalmente	30
hectárea	388	físico	130	dedicar	208	suficientemente	30
control	382	material	129	registrar	208	adecuadamente	29
licor	381	negro	128	comenzar	205	mundialmente	29
tema	375	amplio	127	dejar	204	recientemente	29
cuenta	374	artesanal	126	cosechar	203	independientemente	27
negocio	373	grande/máximo	124	analizar	200	nuevamente	27
árbol	371	determinado	121	mencionar	200	tampoco	27
cacaotero	368	proveniente	121	necesitar	198	claramente	26
centro	364	común	120	formar	196	completamente	26
economía	363	oscuro	119	tostar	196	constantemente	26
base	362	financiero	118	participar	195	económicamente	26

Tableau 6 : 100 premiers lemmes des noms communs, adjectifs, verbes et adverbes du sous-corpus ambiant (TXM)

Source : élaboré par nous



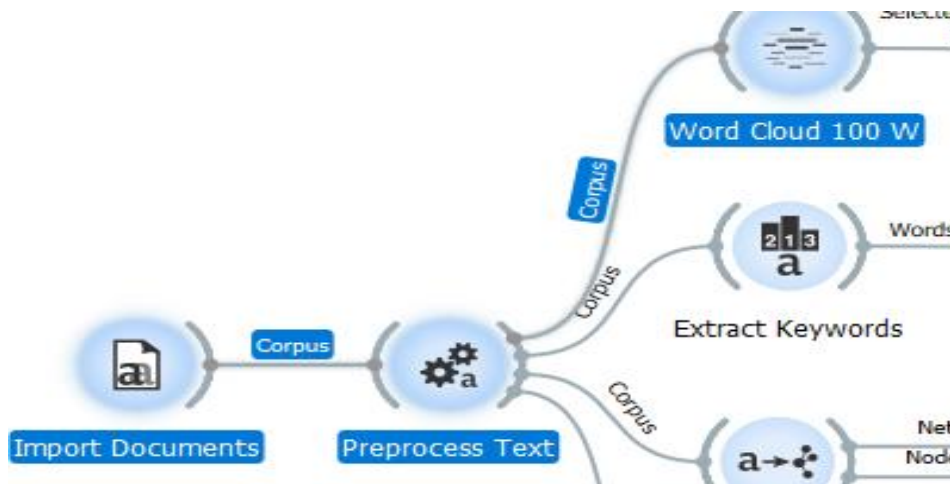


Figure 22 : Flux de travail pour le Nuage de mots du sous-corpus ambiant (Orange)  
 Source : élaboré par nous

Lors de l'importation de documents, le programme accepte une liste d'archives texte au format UTF8 transformée en corpus brut. L'étape de transformation *Preprocess Text* permet d'éliminer les majuscules et les Urls. Elle permet de transformer chaque mot en *token* tout en excluant les numéros et les signes de ponctuation. Finalement, cette étape permet la lemmatisation des *tokens* créés selon le lemmatiseur intégré dans Orange : *UDPipe lemmatizer*<sup>39</sup>. Le sous-corpus ainsi préparé peut alimenter les différentes requêtes disponibles dans le programme comme *Word cloud*. Associé au nuage de mots, la recherche de mots-clés *Extract Keywords* selon la méthodologie YAKE ! (Campos et coll., 2020) appliqué à l'espagnol<sup>40</sup> confirme la spécificité de la thématique abordée, le cacao et le chocolat, mais dans un contexte économique lié à la production nationale et à l'exportation (*Tableau 7* : 100 premiers mots-clés du sous-corpus ambiant). L'écosystème linguistique du sous-corpus ambiant présente dès le début de l'analyse une forte tendance liée à l'économie de marché. L'algorithme YAKE ! présenté par Campos *et coll.* permet l'extraction de mots-clés sans recourir à l'utilisation d'un corpus de référence ce qui permet de l'appliquer à plusieurs langages. L'extraction se fait directement à partir du corpus en cinq étapes qui permettent d'attribuer un score à chaque mot-clé déterminant ainsi le degré de spécificité du mot dans le corpus étudié

<sup>39</sup> [https://ufal.mff.cuni.cz/udpipe/1/users-manual#model\\_training\\_tagger](https://ufal.mff.cuni.cz/udpipe/1/users-manual#model_training_tagger)

<sup>40</sup> Algorithme intégré dans le programme Orange

(2020 : 261-271). Dans le programme Orange, plus le score est élevé, plus le mot est spécifique.

Token	Yake!	Token	Yake!	Token	Yake!	Token	Yake!
ecuatoriano	0,034	guayaquil	0,009	radialistas	0,005	variedad	0,004
mundo	0,029	formulario	0,008	roberto	0,005	ecuatorianos	0,004
aroma	0,028	nivel	0,004	sur	0,005	mejores	0,004
fino	0,027	países	0,004	unesco	0,005	detrás	0,004
ecuador	0,027	corresponde	0,008	empresas	0,005	siglo	0,004
calidad	0,025	válido	0,008	reconocido	0,005	granos	0,004
sabor	0,025	erróneo	0,008	campos	0,005	peralta	0,004
producto	0,022	incorrecta	0,008	marcados	0,005	exportaciones	0,004
país	0,022	trabajo	0,008	obligatorios	0,005	sello	0,004
chocolate	0,020	pepa	0,008	tipo	0,005	leer	0,004
producción	0,018	motivos	0,008	mantener	0,005	chocolateros	0,004
chocolates	0,018	cacaos	0,008	guayas	0,005	salón	0,004
productos	0,016	origen	0,008	asociación	0,005	perfil	0,004
años	0,016	marca	0,007	apasionados	0,005	explosión	0,004
arriba	0,015	quito	0,007	exportadores	0,005	reintentar	0,004
nacional	0,014	internacional	0,007	año	0,005	minutos	0,004
productores	0,013	grano	0,007	parís	0,004	identificador	0,003
noticias	0,013	investigación	0,007	sector	0,004	recuerdo	0,003
cacao	0,012	amazonía	0,006	viaje	0,004	industria	0,003
mercado	0,011	materia	0,006	ríos	0,004	destinadas	0,003
mundial	0,011	oro	0,006	centro	0,004	estrategias	0,003
proceso	0,010	cultivo	0,006	época	0,004	incrementar	0,003
historia	0,009	sabores	0,006	inválido	0,004		
nombre	0,009	favor	0,006	ruta	0,004		
funcionase	0,009	evento	0,006	maestro	0,004		

Tableau 7 : 100 premiers mots-clés du sous-corpus ambient (Orange)

Source : élaboré par nous

Aussi bien les cent noms communs, adjectifs, verbes, et adverbes les plus fréquents, que le nuage de mot, que l'extraction de mots-clés, ces paramètres confirment la spécificité du discours représenté par le sous-corpus ambient. Le principal sujet est la production nationale du cacao et du chocolat comme produit d'exportation de haute qualité. Il est fait référence à l'origine et au type de cacao produit : la province del Oro, de los Ríos,



de Guayas et l'Amazonie ; le cacao type *Nacional* et *Arriba* aussi appelé *pepa de oro*. Dans l'analyse des thèmes spécifiques, le second thème fait référence à l'activité agrotouristique en mentionnant le concept de *ruta* où il est possible de visiter les *haciendas* qui produisent le cacao et de goûter le chocolat que les touristes font eux-mêmes. Les mots relatifs à la production sont suffisamment explicites pour être directement associés à ce domaine : *producto(s), producción, productores, procesos, trabajo, materia, cultivo, empresas, campos, asociación, chocolateros, industria*. Ceux relatifs au commerce extérieur le sont tout autant : *mundo, mercado, países, marca, internacional, evento, exportadores, París, viaje, exportaciones, salón, estrategias*. Les mots qui décrivent le produit s'inscrivent dans un discours plus marketing que sensoriel et sont liés au cacao plus qu'au chocolat : *aroma, fino, calidad, sabor, sabores*. En effet, les concordances de ces termes (*Annexe 11 : Concordances sous-corpus ambiant (TXM)*) orientent vers cette interprétation : *Chocolate vintage, los granos finos o de aroma , las variedades de más alta calidad que se usan en productos ; podemos llenarnos de adjetivos para describir su dulce sabor su inconfundible aroma su delicada textura pero ninguna palabra es capaz de...* Le discours est représentatif d'une nation productrice de cacao qui fait ses débuts dans le marché international de la production de chocolat pour les *apasionados*, seul mot-clé qui fasse référence aux consommateurs.

Afin de déterminer plus en détail les relations entre les mots significatifs et avoir une idée plus précise de la carte sémantique du discours étudié, le programme Orange offre la possibilité de calculer la distance à partir du cosinus de l'angle de deux vecteurs (*Figure 23 : Flux de travail appliqué aux 100 premiers mots et aux mots-clés du sous-corpus ambiant*).

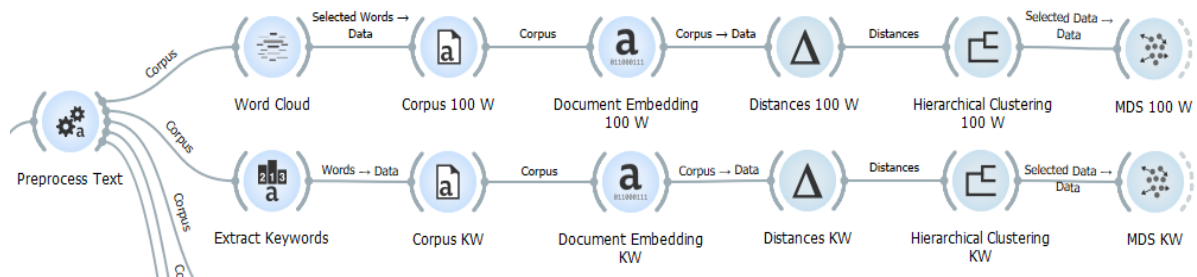


Figure 23 : Flux de travail appliqué aux 100 premiers mots et aux mots-clés du sous-corpus ambiant (Orange)

Source : élaboré par nous

À partir du nuage de mots, le programme Orange permet la création d'un corpus qui, transformé par la fonction *Document Embedding*, permet d'obtenir un vecteur pour chaque tokens à partir d'un modèle pré-entraîné (Grave et coll., 2018 : 14-15). Dans le programme Orange, l'obtention des vecteurs des *tokens* sélectionnés, spécifiques au corpus étudié, est transparente pour le chercheur. Cependant, le calcul se fait à partir d'un corpus entraîné sur Wikipédia (797362600 tokens et 1337109 termes pour l'espagnol) et Common Crawl Corpus (72493240645 tokens et 10614696 termes pour l'espagnol) en appliquant le programme fastText<sup>41</sup>. La description détaillée de cet algorithme n'étant pas l'objet principal de ce travail, nous renvoyons le lecteur aux publications de Bojanowski et coll. (2017) et Graves et coll. (2018) pour de plus amples informations. *Document Embedding* analyse les termes sélectionnés – dans notre cas les 100 premiers mots significatifs et les 100 premiers aux mots-clés – et obtient un vecteur de 300 dimensions pour tous les termes de la liste qui représente la probabilité d'observer dans le corpus étudié les termes généralement associés dans un corpus entraîné (Figure 24 : Exemple de vectorisation des 20 premiers mots les plus fréquents (Orange)) et ainsi mesurer quantitativement la différence du réseau sémantique de chaque terme dans los sous-corpus d'étude avec le réseau sémantique dans lequel ce terme est généralement utilisé.

<sup>41</sup> <https://fasttext.cc/>

Info		embedding-featu	Words	Dim1	Dim2	Dim3	Dim4	Dim5
100 instances (no missing data) 300 features No target variable. 1 meta attribute		hidden	True words	True	True	True	True	True
Variables		title						
<input checked="" type="checkbox"/> Show variable labels (if present)								
<input checked="" type="checkbox"/> Visualize numeric values								
<input checked="" type="checkbox"/> Color by instance classes								
Selection								
<input type="checkbox"/> Select full rows								
1	salir		0.0551764	-0.0704605	0.0410228	0.0710333	-0.00125831	
2	bastante		0.0197377	-0.00610392	0.0397736	0.0173939	-0.026397	
3	probado		0.0157915	0.0263323	0.0045842	0.0237341	-0.0213418	
4	entonces		0.0187523	-0.0382606	-0.0218318	9.84797e-05	0.0224163	
5	hablar		0.0249976	-0.0559955	0.0601083	-0.0169993	-0.0392954	
6	prueba		0.0374416	0.0724606	-0.0142588	0.0782508	0.0213249	
7	gracia		-0.0726775	-0.0183504	0.051136	-0.0147462	0.018803	
8	fino		-0.0867756	-0.114277	0.00657177	-0.0648878	0.0326307	
9	país		0.0926412	-0.0778699	0.0620115	0.000632736	-0.0402646	
10	nacional		0.0277786	-0.00173925	0.0702433	0.0140324	-0.0172833	
11	trabajo		0.0124264	-0.038241	0.0192512	0.0280915	-0.0126495	
12	nivel		-0.029836	-0.0385287	0.0113373	0.0437216	-0.0403488	
13	creer		-0.0227395	-0.0383996	0.0513446	0.0111269	0.0135659	
14	productor		-0.016023	-0.0298766	0.0337462	0.00345886	-0.0181578	
15	pensar		0.022023	-0.055731	0.0338242	-0.0124364	0.0177914	
16	cien		-0.0990221	-4.97699e-06	-0.0384967	-0.0811099	0.0986848	
17	vez		-0.0153777	-0.013886	-0.0811353	0.0792888	0.0282757	
18	ejemplo		0.0260904	-0.0332618	-0.00731317	0.000218015	-0.0160032	
19	osear		0.0274178	0.0363967	0.0959523	0.0421208	-0.0355067	
20	decir		0.0227086	-0.0439867	0.0228092	-0.0176584	-0.0193668	

Figure 24 : Exemple de vectorisation des 20 premiers mots les plus fréquents (Orange)  
Source : élaboré par nous

En effet, en accord avec l'hypothèse distributionnelle de Harris (1954), le vecteur chiffre les différences des probabilités de présence d'un token dans l'environnement du token vectorisé dans le corpus étudié en prenant comme référence les probabilités de présence de ce token dans un corpus entraîné. Par conséquent, le vecteur représente le réseau sémantique spécifique du terme vectorisé dans le corpus étudié en mesurant la différence avec l'usage général de ce terme dans un corpus entraîné. Après avoir obtenu que chaque terme soit représenté par un vecteur, il est possible d'appliquer des opérations mathématiques pour caractériser les différences et similitudes entre ces vecteurs, par conséquent, entre les termes sélectionnés.

Distance Matrix - Orange

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1		0,743	0,836	0,689	0,519	0,781	0,786	0,860	0,731	0,837	0,681	0,809	0,621	0,891	0,564	0,861	0,694	0,764	0,674	0,618
2	0,743		0,681	0,741	0,725	0,753	0,691	0,629	0,850	0,849	0,739	0,728	0,803	0,890	0,718	0,771	0,662	0,732	0,857	0,689
3	0,836	0,681		0,744	0,777	0,596	0,862	0,759	0,893	0,891	0,838	0,854	0,736	0,911	0,826	0,865	0,761	0,790	0,908	0,686
4	0,689	0,741	0,744		0,691	0,759	0,757	0,772	0,660	0,804	0,738	0,766	0,671	0,784	0,603	0,819	0,510	0,505	0,757	0,527
5	0,519	0,725	0,777	0,691		0,798	0,768	0,809	0,808	0,813	0,714	0,778	0,549	0,808	0,408	0,823	0,775	0,645	0,676	0,392
6	0,781	0,753	0,596	0,759	0,798		0,682	0,843	0,939	0,813	0,793	0,753	0,743	0,944	0,772	0,830	0,644	0,737	0,840	0,750
7	0,786	0,691	0,862	0,757	0,768	0,682		0,759	0,931	0,853	0,833	0,823	0,689	0,918	0,746	0,901	0,715	0,781	0,844	0,763
8	0,860	0,629	0,759	0,772	0,809	0,843	0,759		0,818	0,852	0,719	0,757	0,898	0,768	0,839	0,861	0,847	0,756	0,905	0,777
9	0,731	0,850	0,893	0,660	0,808	0,939	0,931	0,818		0,458	0,731	0,679	0,797	0,676	0,793	0,856	0,739	0,742	0,830	0,722
10	0,837	0,849	0,891	0,804	0,813	0,813	0,853	0,852	0,458		0,794	0,539	0,866	0,689	0,868	0,861	0,764	0,831	0,883	0,793
11	0,681	0,739	0,838	0,738	0,714	0,793	0,833	0,719	0,731	0,794		0,708	0,802	0,670	0,705	0,884	0,838	0,666	0,882	0,795
12	0,809	0,728	0,854	0,766	0,778	0,753	0,823	0,757	0,679	0,539	0,708		0,916	0,730	0,857	0,866	0,770	0,703	0,882	0,787
13	0,621	0,803	0,736	0,671	0,549	0,743	0,689	0,898	0,797	0,866	0,802	0,916		0,895	0,294	0,787	0,742	0,782	0,729	0,426
14	0,891	0,890	0,911	0,784	0,808	0,944	0,918	0,768	0,676	0,689	0,670	0,730	0,895		0,888	0,948	0,906	0,787	0,949	0,834
15	0,564	0,718	0,826	0,603	0,408	0,772	0,746	0,839	0,793	0,868	0,705	0,857	0,294	0,888		0,799	0,720	0,650	0,668	0,402
16	0,861	0,771	0,865	0,819	0,823	0,830	0,901	0,861	0,856	0,861	0,884	0,866	0,787	0,948	0,799		0,737	0,881	0,975	0,815
17	0,694	0,662	0,761	0,510	0,775	0,644	0,715	0,847	0,739	0,764	0,838	0,770	0,742	0,906	0,720	0,737		0,668	0,794	0,655
18	0,764	0,732	0,790	0,505	0,645	0,737	0,781	0,756	0,742	0,831	0,666	0,703	0,782	0,787	0,650	0,881	0,668		0,809	0,562
19	0,674	0,857	0,908	0,757	0,676	0,840	0,844	0,905	0,830	0,883	0,882	0,882	0,729	0,949	0,668	0,975	0,794	0,809		0,557
20	0,618	0,689	0,686	0,527	0,392	0,750	0,763	0,777	0,722	0,793	0,795	0,787	0,426	0,834	0,402	0,815	0,655	0,562	0,557	

Figure 25 : Matrice de distance des 20 premiers mots les plus fréquents (Orange)  
 Source : élaboré par nous

Une de ces opérations est le calcul de la distance entre deux vecteurs qui est déterminée par le cosinus de l'angle entre ces deux vecteurs en divisant leur produit vectoriel par le produit de leur norme. Sans détailler les processus algorithmiques de ces calculs statistiques, concrètement, cette distance permet de comparer le cadre sémantique dans lequel est utilisé chaque token sélectionné, elle est nommée « distance sémantique » dans la suite de cette étude. Dans le programme Orange, le calcul de la distance sémantique est transparent. Le module *Distance* retourne une matrice dans laquelle les distances entre chaque mot sont consignées (Figure 25 : Matrice de distance des 20 premiers mots les plus fréquents (Orange)).

Une fois les distances sémantiques obtenues, il reste à appliquer une méthode statistique de comparaison de ces distances pour visualiser les termes qui s'emploient dans un cadre sémantique similaire (Figure 26 : Regroupement hiérarchique des 100 premiers mots du sous-corpus ambiant). Dans cette étude, la méthode de variance minimale de

Ward (Saporta, 2006 : 258) disponible dans la programme Orange a été sélectionnée pour le tracer de la fonction *Hierarchical clustering*, car elle permet un regroupement agglomératif qui regroupe les vecteurs selon leur proximité de façon à identifier les termes qui possèdent un réseau sémantique proche. Le dendrogramme permet la visualisation du regroupement des données selon les distances qui séparent les tokens vectorisés. La largeur des bâtons horizontaux est uniquement déterminée par l'espace vertical entre les termes, leur longueur dépend de la distance sémantique calculée. Par exemple, dans la *Figure 26* : Regroupement hiérarchique des 100 premiers mots du sous-corpus ambiant, ce calcul permet de déterminer que les tokens *sabor* et *aroma* sont sémantiquement les plus proches parmi les 100 premiers termes du sous-corpus ambiant. Ces deux termes forment alors un nœud *sabor/aroma* qui est sémantiquement proche du nœud *grano/fino*. Pour l'analyse, les couleurs du dendrogramme matérialisent de manière arbitraire dix regroupements. Une fois les algorithmes appliqués, les termes réunis dans un même regroupement partagent un même contexte sémantique. Sans qu'ils soient forcément cooccurrents, bien qu'ils puissent l'être, plus leurs cooccurrents présentent des similitudes, plus ils seront proches dans la hiérarchie des regroupements.

Dans le premier regroupement (C1), apparaissent les mots les plus fréquemment associés entre eux dans le corpus : *cacao, chocolate, aroma, sabor, fino, grano*. Il est effectivement difficile de trouver un texte sur le cacao équatorien sans que *fino de aroma, grano* ou *chocolate* n'y soient pas associés. Une phrase prototypique de ce regroupement pourrait être : « *el grano Fino de Aroma es un cacao que produce un chocolate de gran sabor*<sup>42</sup> ». Cette phrase est prototypique du regroupement, elle est proposée comme une possibilité de mise en discours des termes de C1, elle ne fait pas partie du corpus et ne peut donc pas être illustrée par des concordances. Les termes *fino* et *aroma* sont associés pour identifier le type de cacao équatorien : *cacao fino de aroma*. En effet, dans le sous-corpus ambiant, *aroma* présente une fréquence de 1522, *fino* de 1413 et leur cofréquence est de 1069 avec une distance moyenne de 1,5 token entre les

---

<sup>42</sup> La fève Fino de Aroma est un cacao qui produit un chocolat de grande saveur (traduit par nous)

deux termes. Par analogie, cette expression a le même poids sémantique que les cépages dans le domaine du vin, le *fino de aroma* serait considéré comme un cépage noble dont la seule mention garantirait l'exceptionnelle qualité du chocolat produit.

Le deuxième regroupement (C2) est directement lié au processus de production du cacao et à son exportation : *productor, agricultor, orgánico, origen, fermentación, cultivo cosecha, tonelada, exportador, exportación, producción, comercialización*. Dans ce regroupement, il est aussi possible d'identifier la dualité qui existe entre *agricultor* associé à *orgánico, fermentación, cosecha, cultivo* et le *productor* associé à *exportador, exportación, comercialización* et *producción*, ces deux mondes étant étroitement liés par un seul mot : *tonelada* décrivant aussi bien la quantité de cacao produit qu'exporté comme il est possible de le vérifier à partir des concordances en *Annexe II : Concordances sous-corpus ambient (TXM)*. Dans ce regroupement, les grandes étapes de la production du cacao sont représentées : une fois la plante cultivée, le fruit récolté et fermenté, l'agriculteur passe les tonnes de cacao à l'exportateur qui transforme et commercialise les tonnes produites.

Dans le troisième regroupement (C3), la qualité, la variété, la zone géographique et la marque sont les termes les plus souvent utilisés pour décrire le processus de l'entreprise, de l'association ou de l'organisation qui produit du cacao : *calidad, variedad, cadena, marca, provincia, zona, actividad, empresa, organización, asociación*.

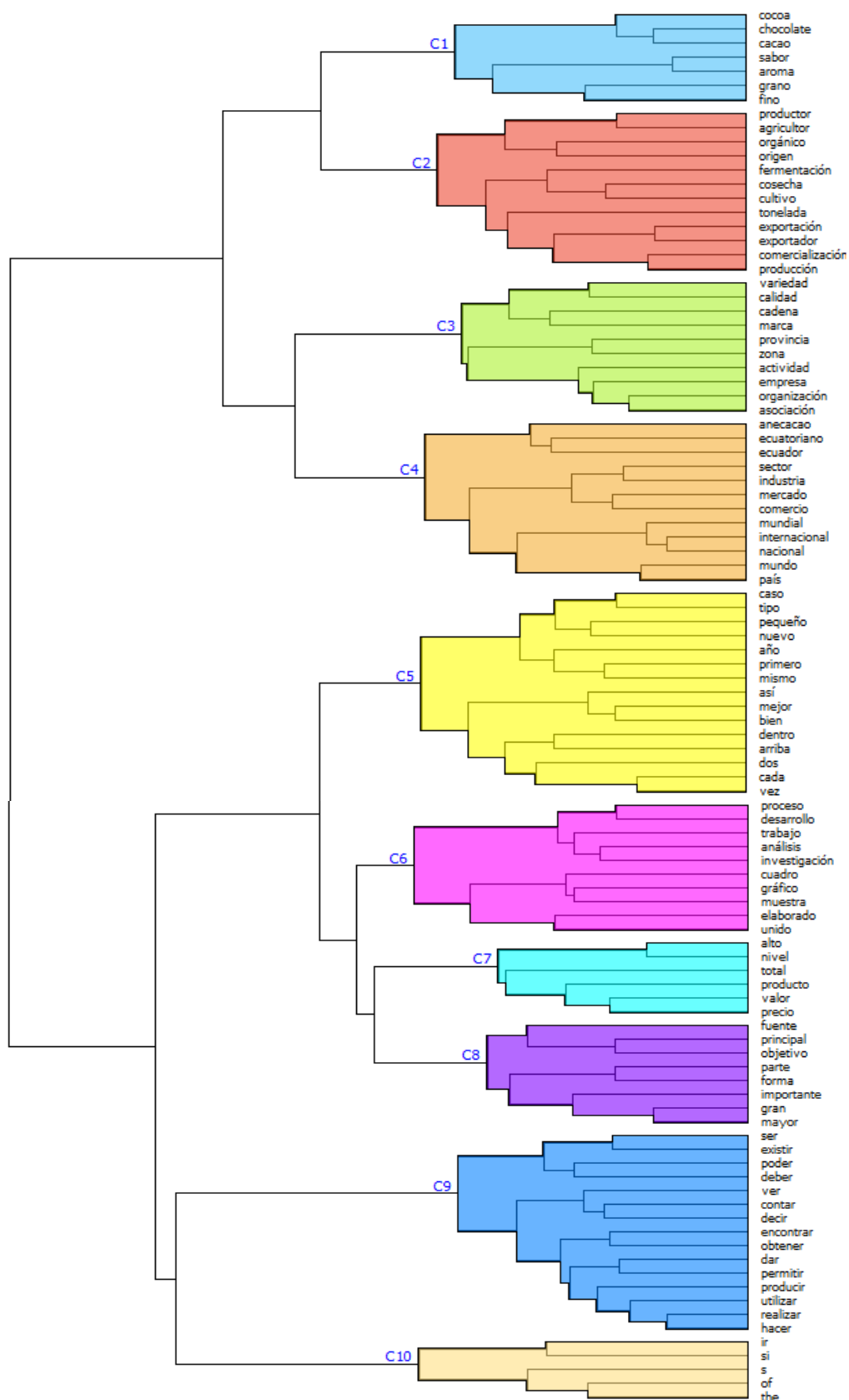


Figure 26 : Regroupement hiérarchique des 100 premiers mots du sous-corpus ambient (Orange)

Source : élaboré par nous

De même le quatrième regroupement (C4) traite des caractéristiques de l'industrie équatorienne du cacao qui possède sa revue officielle : *Anecacao*. Le secteur de cette industrie bien ancré en Équateur est cependant ouvert au commerce et au marché international : *anecacao, ecuador, ecuatoriano, industria, sector, comercio, mercado, mundial, nacional, internacional, país, mundo* (extrait de concordances *Annexe II : Concordances sous-corpus ambient (TXM)*).

Les cinquième et huitième regroupements (C5 et C8) réunissent des adverbes, des adjectifs et des noms neutres et génériques qui n'offrent pas d'informations pertinentes mis à part pour les termes *pequeño* et *arriba*. En effet, l'adjectif *pequeño* est souvent associé à la condition précaire des agriculteurs et des producteurs (concordances *Annexe II : Concordances sous-corpus ambient (TXM)*) qui cultivent le cacao sur de petites surfaces. L'expression *cacao arriba* est la version équatorienne « cépage » *fino de aroma*. Il est possible de trouver des cacaos *fino de aroma* en Colombie, au Costa Rica ou encore au Pérou, tout comme il est possible de trouver du Pinot Noir en Afrique du Sud ou en Argentine, mais le *cacao arriba* est le cacao *fino de aroma* de l'Équateur et sa reconnaissance sur le marché international du chocolat est comparable à la reconnaissance du Pinot Noir de Bourgogne sur le marché international du vin.

Le sixième regroupement (C6) est plus axé sur les activités de recherche : *proceso, desarrollo, trabajo, investigación, análisis, cuadro, gráfico, muestra, elaborado*.

Le septième regroupement (C7) se réfère à la valeur économique du produit : *nivel, alto, total, producto, valor, precio*.

Le neuvième regroupement (C9), mis à part les verbes d'état, contient une liste d'actions le plus souvent liées à la description de processus : *ser, existir, poder, deber, ver, decir, contar, encontrar, obtener, permitir, dar, producir, utilizar, realizar, hacer*.

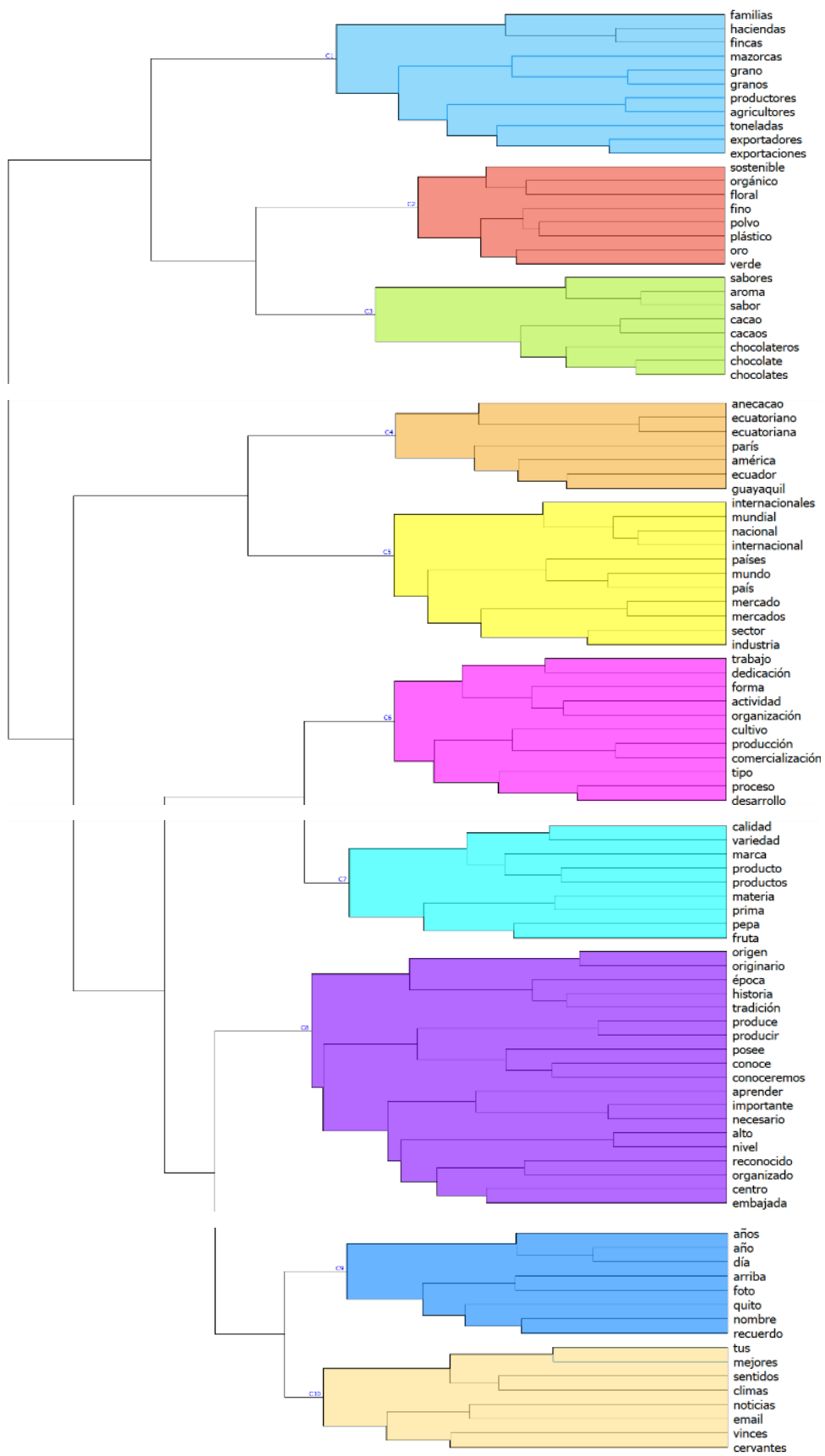
Le dernier regroupement ne présente pas d'intérêt particulier.



Il est aussi possible d'obtenir ce regroupement hiérarchique à partir des mots-clés du sous-corpus en suivant le même processus (*Figure 27 : Regroupement hiérarchique des mots-clés du sous-corpus ambiant*), et ainsi observer les différences de regroupements sémantiques entre les termes d'occurrences les plus élevées et les termes spécifiques du discours considéré. Les regroupements obtenus à partir de deux sources différentes offrent des résultats complémentaires.

Le premier regroupement des mots-clés (C'1) introduit l'environnement de production avec les *haciendas* dont le modèle correspond davantage au concept de monoculture intensif de la production côtière avec des étendues de plusieurs hectares où les producteurs exportent de grandes quantités de cacao. Les *fincas* possèdent moins d'hectares dédiés au cacao et sont associées à une exploitation familiale d'agriculteurs qui cultivent différents produits. En incluant dans ce regroupement les fèves et les cabosses, cela confirme que l'Équateur est avant tout un exportateur de cacao même si dans les deux regroupements suivants il est question du cacao et du chocolat.

Le deuxième et troisième regroupement (C'2 et C'3) traitent des caractéristiques du chocolat et du cacao. Les termes les plus souvent utilisés pour les décrire sont : *sostenible, orgánico, floral, fino, sabor, aroma*. L'adjectif *sostenible* n'est pas une caractéristique intrinsèque du cacao ou du chocolat, mais du processus de production et de l'économie du domaine. L'adjectif *orgánico* est autant une caractéristique du système de production que du produit ; les concordances retournent des expressions comme : *chocolate orgánico, cacao orgánico, de origen orgánico, cultivo orgánico, código orgánico de la economía, abono orgánico, azucar orgánico*, etc. En ce qui concerne l'adjectif *floral*, c'est le seul adjectif du discours sensoriel présent dans la roue des arômes (*Figure 13 : Roues des arômes du chocolat, de l'huile d'olive et du vin*) qui apparaît comme mot-clé. Il est aussi présent dans les 100 premiers adjectifs lemmatisés (*Tableau 6 : 100 premiers lemmes des noms communs, adjectifs, verbes et adverbes du sous-corpus ambiant*) en 59 position.



*Figure 27 : Regroupement hiérarchique des mots-clés du sous-corpus ambient (Orange)*

*Source : élaboré par nous*

Bien que le discours sensoriel n'apparaisse pas dans les cent premiers mots significatifs du sous-corpus ambient, il compte sur au moins un représentant dans les mots-clés. Les concordances de cet adjectif mettent en évidence la présence d'autres adjectifs du discours sensoriel qui ont une fréquence d'utilisation relativement faible devant la taille du sous-corpus (*Annexe 11 : Concordances sous-corpus ambient (TXM)*) : *frutal, nuez, caramelo, malta, amargura, acidez, astringencia, dulce, cítricos, plátano, rosas, grosellas negras, especias, herbal, madera, bosque después de la lluvia, hierba recién cortada, terroso, hongo, musgo, flor de naranja*. Dans le sous-corpus ambient, l'écosystème linguistique dans lequel ces adjectifs sont produits se trouve être celui de dégustations professionnelles faites par des experts appelés *catador*. Comme précédemment commenté, les termes *fino* et *aroma* réapparaissent comme mots-clés pour faire référence au type de cacao équatorien et lui donner ses lettres de noblesse.

Le quatrième regroupement (C'4) peut être considéré avec le cinquième (C'5). Ils sont directement liés et mettent en avant le caractère national et international du marché du cacao équatorien : *anecacao, ecuatoriano, ecuatoriana, parís, américa, ecuador, guayaquil, internacionales, mundial, nacional, internacional, países, mundo, país, mercado, mercados, sector, industria*.

Les sixième (C'6) et septième (C'7) regroupements des mots-clés sont comparables aux sixième (C6) et septième (C7) regroupements des cent premiers mots. Cependant, C'6 ne fait pas référence à la recherche, mais au processus de production et de commercialisation en se focalisant sur l'idée de travail et de dévouement qu'impliquent les activités de culture, de production et de commercialisation de la *pepa* qui apparaît comme la *materia prima* dans C'7. Le regroupement C'7 met en avant le produit sans faire référence à sa valeur économique, mais à sa qualité et sa variété.

En revanche, le huitième regroupement (C'8) ne renvoie pas à un groupe d'adverbe comme le C8, mais à un thème de patrimoine considérant que le cacao est connu

(*conoce, conoceremos*) et reconnu (*reconocido*) dans l'histoire (*historia*) et la tradition (*tradición*) équatorienne comme le confirment les concordances des mots *historia* et *tradición* (*Annexe 11 : Concordances sous-corpus ambient (TXM)*)

Les neuvièmes et dixièmes regroupements font référence à des références spatio-temporelles et des noms propres sans définir exactement un thème particulier : *años, año, día, arriba, foto, quito* (Capitale de l'Équateur), *nombre, recuerdo, tus, mejores, sentidos, climas, noticias, email, vinges* (ville fondée grâce à l'exploitation du cacao en 1879), *cervantes*. Il est intéressant de remarquer que le terme *sentido* apparaît dans le dernier regroupement avec une fréquence de 131. Cependant, les cooncordances mettent en évidence de nombreuses utilisations de l'expression *en este sentido*, du verbe **sentirse**, laissant peu de cas directement liés aux sens.

Les regroupements hiérarchiques offrent une lecture plus détaillée du thème central déterminé par le nuage de mots, l'extraction des mots-clés ou l'analyse des thèmes spécifiques. En appliquant ce regroupement sur les mots-clés, il est possible d'identifier des thèmes liés à l'histoire, à l'agrotourisme. Cependant, les regroupements hiérarchisés ne permettent pas de visualiser quels sont les regroupements centraux et quelles relations ils entretiennent. Cette visualisation est possible, dans le programme Orange, grâce à la fonction d'échelle multidimensionnelle (Borg et coll., 2018), *multidimensional scaling (MDS)*. Cette fonction projette les éléments des regroupements sur un plan tout en conservant leur code de couleur pour faciliter leur identification. La projection est réalisée selon un échelonnement multidimensionnel qui ajuste au mieux les distances entre deux termes dans un espace bidimensionnel. La représentation graphique (*Figure 28 : Agrandissement de la projection bidimensionnelle (MDS) des 100 premiers mots du sous-corpus ambient*) offre une visualisation des regroupements calculés à partir des cent premiers mots significatifs qui interagissent le plus. Les regroupements C2, C3, C9 et C10 couvrent la majorité du réseau sémantique définissant le discours représenté par le sous-corpus ambient comme technique, académique et persuasif si l'on considère que le thème central de ce discours est la promotion du cacao équatorien sur le marché international. En effet, la technicité du

processus de production, la qualité de la chaîne de valeur, et les caractéristiques techniques du cacao produit, le tout soutenu par la recherche et l’information statistique, permettent de répondre aux exigences du cahier des charges de l’exportation. La mise en avant de cette technicité semble être l’argumentation privilégiée pour vanter les mérites du cacao équatorien sur le marché international.

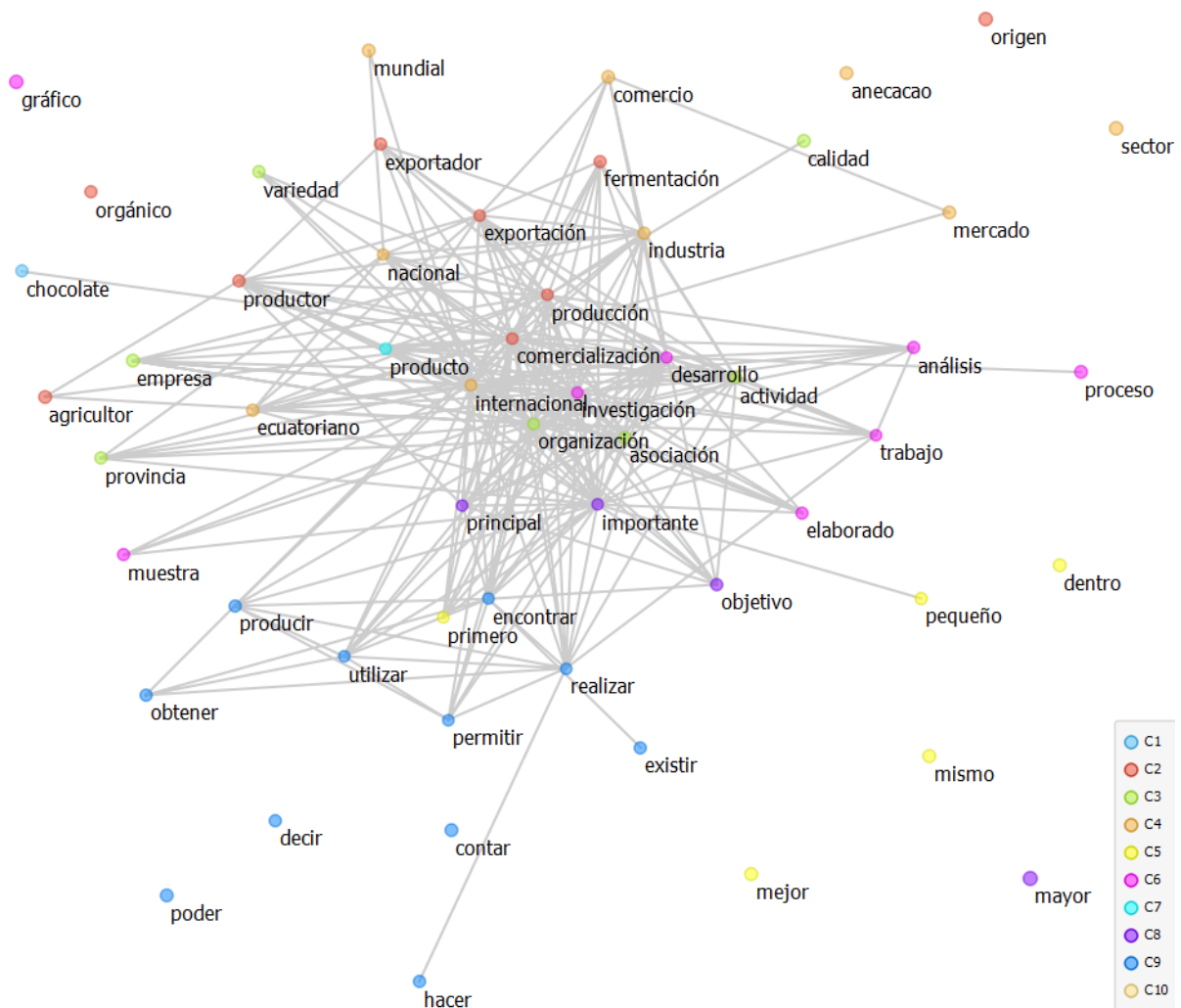


Figure 28 : Agrandissement de la projection bidimensionnelle (MDS) des 100 premiers mots du sous-corpus ambiant (Orange)

Source : élaboré par nous

La représentation graphique de la projection bidimensionnelle (Figure 29 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus ambiant) offre une visualisation des regroupements calculés à partir des mots-clés qui



Les cooccurrences des termes : *cacao* et *chocolate*, obtenue avec le programme TXM, confirment l'analyse précédente (*Tableau 8* : cooccurrence de cacao dans le sous-corpus *ambient*). Les cooccurrences sont présentées selon leur catégorie de partie du discours avec leur fréquence d'utilisation dans le sous-corpus *ambient* (Freq), leur co-fréquence dans une plage de 10 termes à gauche et 10 termes à droite du mot de référence (CoFreq). Leur indice d'affinité (Score) est déterminé à partir du degré de probabilité de cooccurrence lexicale signalant un écart entre co-fréquence réelle et co-fréquence attendue définissant ainsi un indicateur de probabilité de rencontre<sup>43</sup>. Cet indice de spécificité est associé à la distance moyenne exprimée en token entre le cooccurrent et le terme de référence (MDist).

Nom	Freq	CoFreq	Score	MDist	Nom	Freq	CoFreq	Score	MDist
aroma	1522	1068	1000	3,2	exportador	588	203	13	3,5
manteca	472	408	202	3,1	región	392	148	13	4,8
grano	2813	1315	201	2,6	retos	23	21	12	3,0
producción	3001	1201	97	3,6	industria	682	222	11	4,7
exportación	1875	799	97	3,3	edulcorante	33	26	11	8,0
licor	381	255	81	3,6	elaboración	461	161	11	4,5
cultivo	1010	476	75	3,2	correo-e	26	22	11	6,9
pasta	221	169	68	3,3	comerciante	104	53	10	3,6
polvo	318	202	59	3,1	café	233	93	10	3,2
variedad	829	369	51	3,9	<b>Adjectif</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
productor	1803	658	50	3,9	fino	1413	1141	1000	1,0
mundo	858	374	49	4,3	ecuatoriano	1264	649	123	1,9
productor	631	288	42	3,1	mundial	1028	503	86	3,7
tipo	931	379	41	3,5	elaborado	474	292	80	3,0
calidad	2151	721	40	4,2	derivado	145	111	45	2,8
azúcar	294	154	31	4,2	beneficiado	148	107	39	1,2
cadena	887	339	30	3,4	nacional	1023	393	36	3,0
plantación	435	199	29	3,2	derivado	164	106	32	2,4
comercialización	637	260	28	3,6	principal	835	320	29	5,0
baba	91	69	28	2,9	orgánico	788	295	25	2,7
sabor	1380	462	26	4,8	bueno mejor	633	239	21	3,2
torta	109	75	25	4,6	ordinario	74	49	16	2,2

<sup>43</sup> Pour plus de précision sur le calcul de cet indice de spécificité, voir le manuel d'utilisation de TXM : <https://txm.sourceforge.net/doc/manual/manual59.xhtml#toc262>

precio	1221	407	22	4,2	contacto	21	21	14	2,6
gráfico	231	115	21	5,5	seco	318	126	13	3,9
chocolate	3928	1076	20	5,1	industrializado	55	37	12	2,1
números	30	30	20	2,2	convencional	61	39	12	1,6
tonelada	577	220	20	4,9	especial	280	111	11	3,9
hectárea	388	162	19	4,2	semielaborado	37	28	11	4,6
semielaborados	84	57	19	4,1	sensorial	112	56	10	4,2
partida	147	81	18	5,3	exportado	217	89	10	4,0
destino	264	121	18	4,4	proveniente	121	58	10	2,9
demanda	524	197	17	4,1	destino	15	15	10	4,2
país	2640	740	16	5,5	<b>Verbe</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
árbol	371	148	15	3,8	producir	504	227	32	3,2
pasta	38	32	15	1,4	ser	10636	2736	30	4,6
ingrediente	222	100	14	4,4	cultivar	119	75	22	2,9
adición	55	39	14	3,7	exportar	206	102	18	4,5
industrialización	100	57	14	3,5	corresponder	210	97	15	4,5
procesamiento	174	82	13	3,7	procesar	46	33	12	3,2
porcentaje	306	123	13	4,9	sembrar	264	103	10	4,4

Tableau 8 : cooccurrence de cacao dans le sous-corpus ambiant (TXM)<sup>44</sup>

Source : élaboré par nous

Mis à part le terme *sensorial* présent comme cooccurrent de *cacao*, le discours sensoriel semble être absent de ce sous-corpus. Les adjectifs cooccurrents de *cacao* sont descriptifs occultant toute dimension hédonique ou émotive. L'adjectif *fino* s'utilise pour définir le type de cacao : *fino de aroma*, *ecuatoriano* et *nacional* défini l'origine. L'adjectif *sensorial* présente 30 cooccurrences dans lesquelles il apparaît être plus associé à la qualité du cacao qu'à un discours sensoriel : *procesos de fermentación de cacao para maximizar el perfil sensorial*, *factores post-cosecha sobre la calidad sensorial del cacao criollo*, *el perfil sensorial de sabores del cacao*, *influye directamente en la calidad física y sensorial del cacao*, *para disminuir la acidez y mejorar la calidad sensorial del cacao*, *protocolo de evaluación sensorial en cacao*, *experto en evaluación sensorial*. La sensorialité du discours ambiant se réfère plus à des

<sup>44</sup> Freq. signifie fréquence du mot ; Cofreq signifie fréquence de l'association mot + cacao ; le Score représente l'indice de spécificité qui détermine la force du lien avec *cacao* ; MDist est la distance moyenne en mot entre le terme considéré et *cacao*.



critères analytiques guidés par un profil qui rappelle la roue des arômes, qu'à des critères émotifs.

Les cooccurrents du terme *chocolate* semblent être tout aussi descriptifs : *negro, ecuatoriano, oscuro, artesanal, caliente, puro*, définissent les caractéristiques physiques et l'origine du chocolat. Cependant, des termes comme *delicioso, amargo, puro*, font leur apparition, mais ils sont loin d'être des représentants significatifs d'un discours sensoriel ou même émotif. En effet, la lecture des cooccurrences met en évidence une utilisation de *delicioso* dans un discours plus promotionnel que sensoriel : *oscuramente delicioso el chocolate ecuatoriano contrasta con la inocencia del cremoso interior de vainilla, el delicioso sabor de un exquisito chocolate elaborado a base de cacao ecuatoriano, el delicioso y mundialmente conocido chocolate, oscuro saludable y delicioso con Pacari se reinventa un nuevo tipo de chocolate, un chocolate orgánico delicioso bueno para ti*. Les cooccurrents les plus représentatifs décrivent les différentes formes du chocolat : *barra, licor, bombón, polvo, tableta, pasta, bebida* ; et les ingrédients : *leche, macadamia, cacao, ingrediente, azúcar* ; ou encore une description d'aspect physique : *negro, oscuro, caliente*. Les verbes sont liés au processus d'élaboration : *elaborar, producir, templar, amantar, tostar, moldar*. Il est intéressant de noter que la première marque qui apparaît comme cooccurrent est équatorienne ; *pacari*. Ce chocolat a obtenu plusieurs prix internationaux, il est devenu une référence pour présenter le chocolat équatorien sur le marché international (Tableau 9 : cooccurrence de chocolate dans le sous-corpus ambiant (TXM)).

Noms	Freq	CoFreq	Score	MDis	Adjectifs	Freq	CoFreq	Score	MDis
barra	423	217	161	2,2	negro	128	71	56	1,0
sabor	1380	261	75	4,7	ecuatoriano	1264	175	32	2,7
chocolate	3928	460	61	6,0	oscuro	119	50	32	2,2
leche	225	92	57	3,6	bueno / mejor	633	100	23	3,2
confitería	85	57	52	4,2	artesanal	126	36	17	3,7
fabricante	163	64	39	1,5	delicioso	68	25	15	2,4
mundo	858	147	37	4,7	final	295	50	13	4,9
macadamia	98	50	37	2,2	amargo	77	24	12	3,8

licor	381	94	37	4,6	caliente	43	18	12	1,7
consumo	494	106	36	3,1	alemán	63	21	11	5,8
fabricación	134	55	35	2,4	luisa	22	13	11	5,8
bombón	83	41	30	3,0	puro	65	20	10	3,9
elaboración	461	88	26	3,2	<b>Verbes</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
producto	3609	331	24	4,8	ser	10636	795	28	4,6
industria	682	102	21	3,5	elaborar	474	91	27	3,4
cacao	13014	891	20	5,1	producir	239	46	14	2,3
ingrediente	222	52	20	5,3	consumir	56	22	14	2,2
marca	595	90	19	3,8	hacer	142	34	13	3,8
polvo	318	61	18	5,0	templar	25	15	13	3,7
tableta	24	17	16	3,3	buscar	23	14	12	4,4
taza	25	17	16	2,6	amantar	11	10	12	2,5
edulcorantes	19	15	16	1,4	tener	2473	206	11	5,1
pasta	221	46	15	4,2	tostar	188	36	11	6,3
mezcla	200	41	13	5,0	hacer	320	47	10	2,6
azúcar	294	48	12	5,5	moldar	22	12	10	4,3
tabletas	90	25	11	3,8	<b>Noms Propres</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
consumidor	476	61	10	4,4	pacari	136	33	13	4,1
bebida	110	26	10	3,2	europa	275	43	10	4,9
receta	90	23	10	5,7					
aroma	1522	136	10	4,9					

Tableau 9 : cooccurrence de chocolate dans le sous-corpus ambiant (TXM)<sup>45</sup>

Source : élaboré par nous

Les tableaux de cooccurrences permettent d'obtenir une description du réseau sémantique d'un terme donné. Le programme Orange, avec les outils *Corpus to Network* et *Network Explorer*, permet d'obtenir une représentation graphique du réseau sémantique de plusieurs termes sélectionnés selon trois paramètres : la fréquence

<sup>45</sup> Freq. signifie fréquence du mot ; Cofreq signifie fréquence de l'association mot + cacao ; le Score représente l'indice de spécificité qui détermine la force du lien avec cacao ; MDist est la distance moyenne en mot entre le terme considéré et cacao.

absolue, la taille de la fenêtre, et la fréquence de cooccurrence. Avec ces paramètres, il est possible d'obtenir une représentation des cooccurrences dont la fréquence de cooccurrence est supérieure à 200<sup>46</sup>, dans une fenêtre de deux mots à gauche et à droite, des termes dont la fréquence est supérieure à 500 (Figure 30 : Réseau sémantique des principales cooccurrences du sous-corpus ambient).

La représentation graphique confirme les résultats observés dans les tableaux et met en relation plusieurs termes : *cacao fino de aroma*, *pequeño productor*, *producto de alta calidad*, *al nivel mundial*, *el sector cacaotero*. Elle permet de visualiser l'affinité d'un terme : même si *producción* est cooccurent de *chocolate*, le mot est plus souvent associé à *cacao* ; alors que *producto de alta calidad* s'adresse aussi bien au *cacao* qu'au *chocolate*.

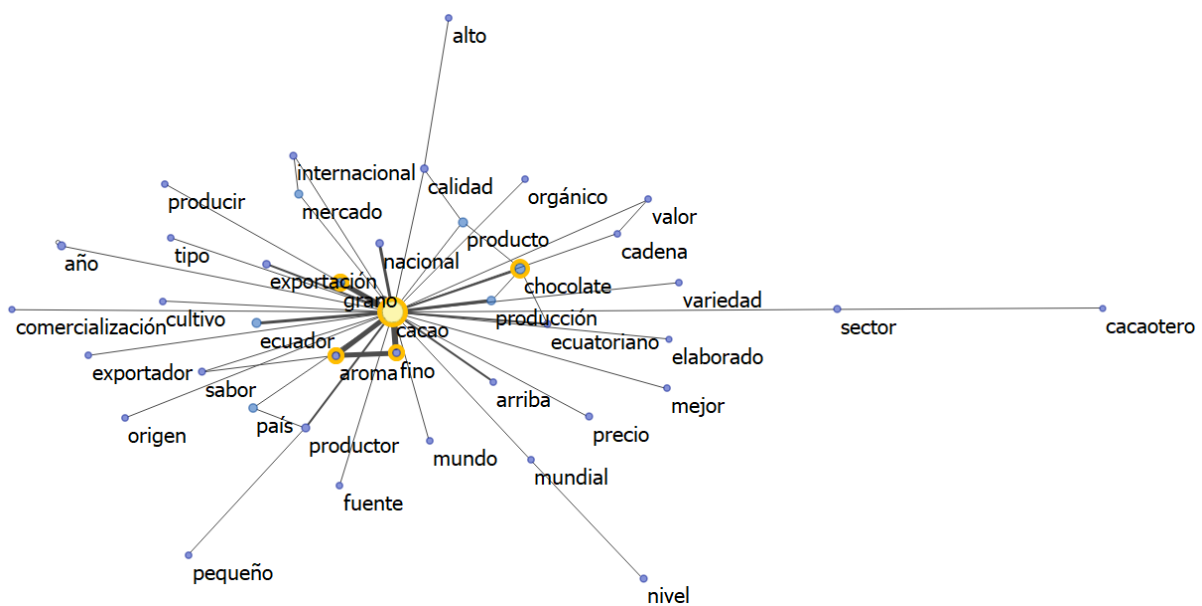


Figure 30 : Réseau sémantique des principales cooccurrences du sous-corpus ambient (Orange)

Source : élaboré par nous

<sup>46</sup> En baissant le paramètre de la fréquence de cooccurrence à 100, les relations sont plus détaillées, mais la représentation graphique devient illisible.

Malgré l'absence remarquée de termes hédoniques ou encore d'expressions sensorielles, et grâce aux différents outils d'étude disponibles dans le programme Orange, il a été mené une analyse des sentiments (*Sentiment Analysis*) qui prédit le sentiment pour chaque document d'un corpus à partir du lexique multilingue du Data Science Lab<sup>47</sup> comme référence pour déterminer la tendance positive ou négative d'un texte (*Figure 31* : Représentation de l'analyse de sentiment des textes du sous-corpus ambiant). Une valence pondérée positive ou négative est attribuée aux termes du lexique de référence intégré au programme Orange de façon à comparer les termes du corpus étudié. Le score final représente la différence entre la somme des termes positifs et la somme des termes négatifs, normalisée par la longueur du document et multipliée par 100. Le score final reflète le pourcentage de différence de sentiment dans le document par rapport au lexique de référence. En évitant de prendre les extrêmes comme exemple, voici un extrait du texte évalué à 12,25 :

*Es nuestra misión producir de forma sostenible el más auténtico chocolate latinoamericano ecuator, la cuna del cacao. Los climas ecuatoriales únicos del ecuador proveen el ambiente perfecto para cosechar el cacao más fino del mundo - cacao arriba. el cacao ha sido venerado en latinoamérica por más de 5000 años más del 85% del cacao fino de aroma del mundo viene de latinoamerica. por eso lo llamamos "la pepa de oro" conoce más donde empezó nuestro viaje. En 2005, nuestros fundadores ecuatorianos visitaron una boutique de chocolate en Paris, aquí pidieron probar el mejor chocolate del mundo. Para su sorpresa, les dijeron que el cacao para producir este chocolate venía de Ecuador. Este momento dio inicio a nuestro viaje y la misión de producir el chocolate más fino. Lo mejor de dos mundos nuestro chocolate es producido localmente utilizando únicamente ingredientes latinoamericano. Combinamos estas increíbles materias primas con la tecnología y el conocimiento más avanzado, que hemos adquirido de los*

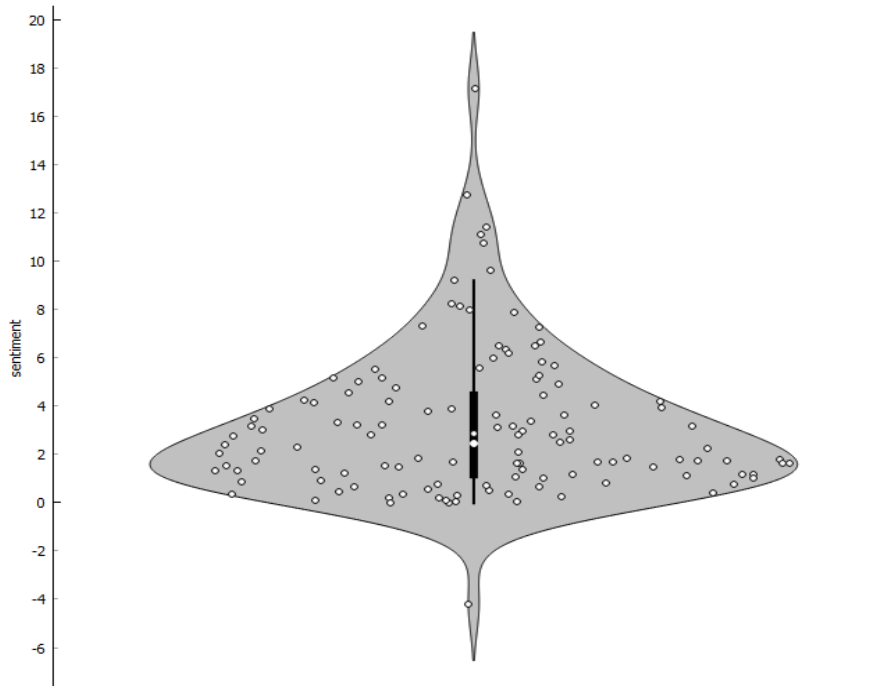
---

<sup>47</sup> <https://orangedatamining.com/widget-catalog/text-mining/sentimentanalysis/>

*chocolateros franceses más finos. Nos gusta pensar en esto como tener lo mejor de dos mundos.*

Voici un extrait du texte évalué -0,02:

*3.3 grano defectuoso. se considera como grano defectuoso a los que a continuación se describen: 3.3.1 grano mohoso. grano que ha sufrido deterioro parcial o total en su estructura interna debido a la acción de hongos, determinado mediante prueba de corte. 3.3.2 grano dañado por insectos. grano que ha sufrido deterioro en su estructura (perforaciones, picados, etc.) debido a la acción de insectos. 3.3.3 grano vulnerado. grano que ha sufrido deterioro evidente en su estructura por el proceso de germinación, o por la acción mecánica durante el beneficiado. 3.3.4 grano múltiple o pelota. es la unión de dos o más granos por restos de mucílago. 3.3.5 grano negro. es el grano que se produce por mal manejo poscosecha o en asocio con enfermedades. 3.3.6 grano ahumado. grano con olor o sabor a humo o que muestra signos de contaminación por humo. 3.3.7 grano plano vano o granza. es un grano cuyos cotiledones se han atrofiado hasta tal punto que cortando la semilla no es posible obtener una superficie de cotiledón.*



*Figure 31 : Représentation de l'analyse de sentiment des textes du sous-corpus ambiant (Orange)*

*Source : élaboré par nous*

Avec une moyenne de deux sur l'échelle des sentiments (ordonnée du point blanc situé sur l'axe central de la *Figure 31 : Représentation de l'analyse de sentiment des textes du sous-corpus ambiant*) et 50% des textes entre un et quatre (selon la barre centrale qui représente les percentiles entre 25 et 75). L'analyse de sentiment des textes du sous-corpus (représentés par les points de la *Figure 31 : Représentation de l'analyse de sentiment des textes du sous-corpus ambiant*) est considérée plutôt positive ce qui n'est pas en contradiction avec un discours dont l'objectif est de faire la promotion du cacao et du chocolat équatorien. Les textes présentant une valence neutre se caractérisent par une objectivité propre aux textes académiques comme le démontre l'extrait précédent.

### 3.2 Dualisme dans le corpus professionnel selon l'activité, l'intention de communication et la langue utilisée

Le sous-corpus constitué à partir des entrevues réalisées aux professionnels du secteur cacao et chocolatier de l'Équateur présente plusieurs écosystèmes linguistiques différents : le professionnel qui produit du chocolat dans un contexte socio-économique équatorien, le professionnel qui produit du cacao sur le sol équatorien lequel évolue aussi bien dans une écolinguistique espagnole imposée que dans une écolinguistique kichwa culturelle. Afin de pouvoir respecter les paramètres du *tertium comparationis* considéré (*Figure 15* : Paramètre du tertium comparationis, le même flux d'analyse est appliqué à chaque sous-corpus Entrevues Professionnelles (*Figure 20* : Flux d'analyse des sous-corpus (Orange). À ce stade de l'analyse, il est important d'avoir une vision claire des différents écosystèmes linguistiques présentés dans la première partie de cette étude. Selon le positionnement théorique de la recherche, le discours de chaque professionnel n'est pas seulement l'empreinte d'une écolinguistique particulière, il en est aussi un des constituants principaux interagissant en accord avec les autres éléments extralinguistiques qui construisent cet environnement écolinguistique. Il a donc été choisi d'analyser séparément le discours du producteur de chocolat, le discours du producteur de cacao quand il s'exprime en espagnol, et son discours quand il s'exprime en kichwa ce qui permet d'obtenir trois moules discursifs propres à trois écosystèmes linguistiques.

### 3.2.1. Moule discursif représentatif du discours professionnel du producteur de chocolat

Les mêmes outils de textométrie ont donc été utilisés pour obtenir les données quantitatives de l'écosystème linguistique dans lequel a été produit le discours professionnel des producteurs de chocolat à commencer par le calcul de la densité lexicale à partir de mille mots aléatoires obtenus par la fonction *shuffle words* de l'outil

*text converter*<sup>48</sup> (Tableau 10 : Paramètres du sous-corpus professionnel producteur de chocolat).

<b>sous-corpus chocolat</b>	<b>SCP</b>	<i>Token</i>	<b>Vocabulaire</b>	<b>Densité lexicale sur 1000 tokens aléatoires</b>
Nombre total		30934	3217	0.358

Tableau 10 : Paramètres du sous-corpus professionnel producteur de chocolat  
Source : élaboré par nous

Malgré la différence de *token* dans la constitution des sous-corpus, la sélection aléatoire de 1000 mots permet la comparaison directe des densités lexicales. L'écart observé avec le sous-corpus ambiant peut, d'une part, s'expliquer par les conditions de production du discours : un écrit produit de manière asynchrone, sans contraintes temporelles, autorise un processus d'édition qui permet l'utilisation d'un vocabulaire plus étendu et plus recherché évitant les répétitions et permettant les améliorations ou les corrections avant la publication. D'autre part, la diversité, la technicité et l'expertise des thèmes abordés dans le discours peuvent aussi avoir une influence directe sur la densité lexicale. L'oralité ne permet pas toujours d'approfondir un thème de manière aussi détaillée que l'écriture où le choix d'une terminologie spécialisée détermine souvent le degré d'expertise du locuteur. L'écologie linguistique de l'oralité, s'inscrivant dans une limite temporelle de synchronicité, n'offre pas toujours au locuteur le temps d'accéder à cette terminologie spécifique de manière posée et réfléchie. Il doit s'y référer spontanément sans interrompre le flux de production du message, expliquant ainsi la différence de densité lexicale observée entre le sous-corpus ambiant et le sous-corpus professionnel producteurs de chocolat ou de cacao. Pour plus de détail sur les différences lexicales entre textes oraux et écrits se référer aux travaux de Nation sur les différentes couvertures lexicales nécessaires pour la compréhension de textes écrits et oraux dans l'apprentissage de l'anglais comme langue étrangère (2006), ou encore les travaux de Fernández sur la détermination de type de texte à partir de la densité lexicale (2018).

<sup>48</sup> <https://textconverter.com/shuffle-words>



En raison de l'extension plus restreinte du sous-corpus professionnel producteur de chocolat que le sous-corpus ambiant, au lieu des cent premiers lemmes des noms, adjectifs, verbes et adverbes, le *Tableau 11* : Lemmes du sous-corpus professionnel producteur de chocolat Tableau 6 : 100 premiers lemmes des noms communs, adjectifs, verbes et adverbes du sous-corpus ambiant réunit les lemmes dont la fréquence est supérieure à dix tout en gardant une classification selon la partie du discours à laquelle ils correspondent.

En première approximation, le résultat de la liste des noms communs les plus utilisés met en avant que le chocolat est plus souvent prononcé que le cacao inversant ainsi la priorité du thème par rapport au sous-corpus ambiant, sans pour autant négliger le cacao qui reste en deuxième position. Cette première constatation n'a rien de surprenante dans un discours prononcé par des producteurs de chocolat qui vantent les mérites et les qualités de la matière première qu'ils utilisent.

<b>[espos="NC"] : eslemma / F</b>							
chocolate	372	tiempo	39	negocio	20	uvilla	14
cacao	324	proceso	37	agricultor	19	idea	13
entrevistador	312	productor	36	familia	19	investigación	13
producto	256	café	33	mj	19	azúcar	12
sabor	103	pero	33	mujer	19	base	12
año	78	producción	33	prueba	19	hierba	12
aroma	70	comunidad	32	zona	19	mes	12
empresa	70	calidad	31	boca	18	premio	12
nombre	62	caso	30	día	18	agregado	11
sensación	61	país	30	colaboración	17	finca	11
persona	59	proyecto	30	valor	17	generación	11
característica	57	ejemplo	29	origen	16	gusto	11
trabajo	54	mercado	29	planta	16	polvo	11
tema	50	mundo	28	fruta	15	porcentaje	11
comercialización	46	nivel	28	productor	15	pregunta	11
palabra	46	responsabilidad	28	leche	15	atención	10
parte	45	vez	27	cliente	14	certificación	10
vino	45	variedad	26	consumidor	14	cultivo	10
gracia	43	manera	22	desarrollo	14	cultura	10

gente	41	provincia	21	em	14	historia	10
marca	41	asociación	20	fermentación	14	paladar	10
barra	40	cosa	20	señor	14	sal	10
tipo	40	momento	20	tarde	14	vainilla	10
<b>[espos="ADJ"]:</b> eslemma / F							
bueno	166	rico	28	solo	17	fuerte	11
nuestro	61	mejor	22	pequeño	15	internacional	11
fino	47	especial	22	frutal	15	mundial	11
diferente	46	grande	19	nuevo	14	agrícola	10
mismo	31	propio	19	ecuatoriano	12	general	10
nacional	30	claro	18	importante	12	natural	10
mucho	29	interesante	18	alto	11		
orgánico	29	perfecto	17	floral	11		
<b>[espos="V.*"]:</b> eslemma / F							
ser	1016	vendar vender	45	recordar	23	salir	13
entrevistar	535	trabajar	43	tomar	22	seguir	13
tener	389	hablar	40	buscar	20	desarrollar	12
haber	254	querer	38	tratar	20	evocar	12
decir	212	aficionar	37	comprar	19	parecer	12
hacer	204	crear creer creer	37	empacar	18	generar	11
estar	197	osear	37	producir	18	mantener	11
probar	130	destacar	35	deber	17	poner	11
poder	93	sentar sentir	33	encontrar	16	quedar	11
ir	86	pensar	31	mirar	16	representar	11
dedicar	67	llamar	28	pasar	16	existir	10
venir	67	saber	27	elaborar	15	mejorar	10
gustar	63	conocer	25	aprovechar	14	sacar	10
dar	52	encantar	25	depender	14		
utilizar	52	llevar	25	ganar	14		
ver	50	empezar	24	comenzar	13		
<b>[espos="ADV"]:</b> eslemma / F							
bien	210	aquí	58	dentro	16	acá	11
más	180	mucho	50	antes	15	actualmente	11
entonces	169	ahí	35	cuándo	14	además	11
ya	149	siempre	35	menos	14	tan	11
también	129	donde	33	obvio	14	tanto	10
muy	93	bastante	24	primero	14		
como	70	así	22	incluso	13		
ahora	59	real	18	poco	13		

*Tableau 11 : Lemmes du sous-corpus professionnel producteur de chocolat dont la fréquence est supérieure à 10 (TXM)*

*Source : élaboré par nous*

Cependant, mise à part la référence faite au vin et au café, les noms communs semblent rester dans une thématique de processus de production et de commercialisation, deux noms qui font partie de la liste. En ce qui concerne le processus de production, les noms des ingrédients sont plus nombreux : *planta, fruta, leche, uvilla, azúcar, hierba, sal, vainilla*. Il est aussi possible d'observer une référence plus directe aux sens, sans pour autant proposer des descriptifs liés aux sensations ressenties : *sabor, aroma, boca, gusto, paladar*.

Les adjectifs les plus fréquents ne semblent pas appartenir à une terminologie spécifique à l'écosystème linguistique du chocolat ou même du cacao. Mis à part *frutal* et *floral* qui s'insèrent dans une thématique liée aux sens, ils ne présentent pas un degré de spécificité important. Ils semblent être des adjectifs courants utilisés dans un écosystème linguistique spécifique pour décrire l'origine d'un grand chocolat organique, naturel et international. Même si c'est un adjectif possessif, la fréquence d'utilisation élevée de *nuestro* démontre un sentiment d'appartenance à un patrimoine commun, soulignant le caractère équatorien du produit tout comme le font les adjectifs : *nacional, propio et ecuatoriano*.

Il en est de même pour les verbes et les adverbes. Cependant, dans la liste des verbes, il existe de plus nombreuses références liées au ressenti laissant croire que la stratégie de communication du discours sensoriel est plus orientée sur l'action que sur la description : *probar, gustar, destacar, sentir, encantar, recordar, evocar, parecer, representar*.

Le nuage de mots, obtenu selon le même flux de travail dans le programme Orange, de la *Figure 32* : Nuage des cent premiers lemmes du sous-corpus professionnel producteur de chocolat réunit les cent lemmes les plus fréquents, indépendamment de leur catégorie

grammaticale, où *chocolate* reste central à côté de *cacao* et *producto*. Les verbes liés à la *sensación* apparaissent dans un deuxième plan, illustrant ainsi l'analyse précédente.



Figure 32 : Nuage des cent premiers lemmes du sous-corpus professionnel producteur de chocolat (Orange)  
 Source : élaboré par nous

Associé aux mots les plus fréquents, l'extraction des cent premiers mots-clés, selon la méthodologie YAKE ! (Campos et coll., 2020) appliqué à l'espagnol<sup>49</sup>, permet de mettre en valeur les termes qui ont une fréquence d'utilisation particulièrement élevée dans cet écosystème linguistique. Un des résultats les plus évidents est la confirmation de l'oralité du discours par la présence d'onomatopées qui le ponctuent : *jajaja*, *mmm*, et la structure de l'entrevue par les mots présents dans les questions posées (surlignés dans le Tableau 12 : 100 premiers mots-clés du sous-corpus professionnel producteur de chocolat).

<b>Word</b>	<b>Yake</b>	<b>Word</b>	<b>Yake</b>	<b>Word</b>	<b>Yake</b>	<b>Word</b>	<b>Yake</b>
-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------

<sup>49</sup> Algorithme intégré dans le programme Orange

nombre	0,10	soy	0,02	talag	0,01	natural	0,01
cuál	0,10	bastante	0,02	probado	0,01	mmm	0,01
cuáles	0,08	sabores	0,02	calidad	0,01	sociedad	0,01
cacao	0,05	tiempo	0,02	tardes	0,01	guayusa	0,01
años	0,05	asociación	0,02	producción	0,01	digamos	0,01
empresa	0,05	sensación	0,02	señor	0,01	determinar	0,01
días	0,05	rico	0,02	chakra	0,01	diversidad	0,01
producto	0,04	jajaja	0,02	señorita	0,01	tostado	0,01
productos	0,04	barras	0,01	tema	0,01	polvo	0,01
comercialización	0,04	marca	0,01	características	0,01	ingresos	0,01
trabajo	0,03	gracias	0,01	mundo	0,01	país	0,01
chocolate	0,03	kallari	0,01	nivel	0,01	específicos	0,01
comunidad	0,03	gourmets	0,01	tipo	0,01	proyecto	0,01
aroma	0,03	laborando	0,01	uvilla	0,01	característica	0,01
personas	0,03	aromas	0,01	tipos	0,01	doce	0,01
gusta	0,03	ingreso	0,01	pensar	0,01	origen	0,01
chocolates	0,02	fino	0,01	energía	0,01	ambiental	0,01
mercado	0,02	asistente	0,01	floral	0,01	conservación	0,01
café	0,02	contable	0,01	prueban	0,01	mami	0,01
responsabilidades	0,02	entendido	0,01	provincia	0,01	bombonería	0,01
sabor	0,02	mes	0,01	misael	0,01	bondades	0,01
dedica	0,02	vendemos	0,01	señora	0,01	artesanal	0,01
viene	0,02	proceso	0,01	vienen	0,01	nacional	0,01
ciento	0,02	gente	0,01	vainilla	0,01	encanta	0,01
ecuador	0,02	shandia	0,01	asocian	0,01	comunidades	0,01

Tableau 12 : 100 premiers mots-clés du sous-corpus professionnel producteur de chocolat (Orange)

Source : élaboré par nous

Il est aussi possible d'observer le nom de fruits comme : *uvilla*, *vainilla*, *guayusa*. *Frutal* n'est pas considéré comme un mot-clé contrairement à *floral*. Dans son ensemble l'analyse des mots-clés apporte des résultats similaires à l'analyse appliquée aux mots qui ont une fréquence supérieure à 10. Cependant, des noms propres apparaissent comme le nom d'une entreprise de production de chocolat : *Kallari* ; le nom d'un

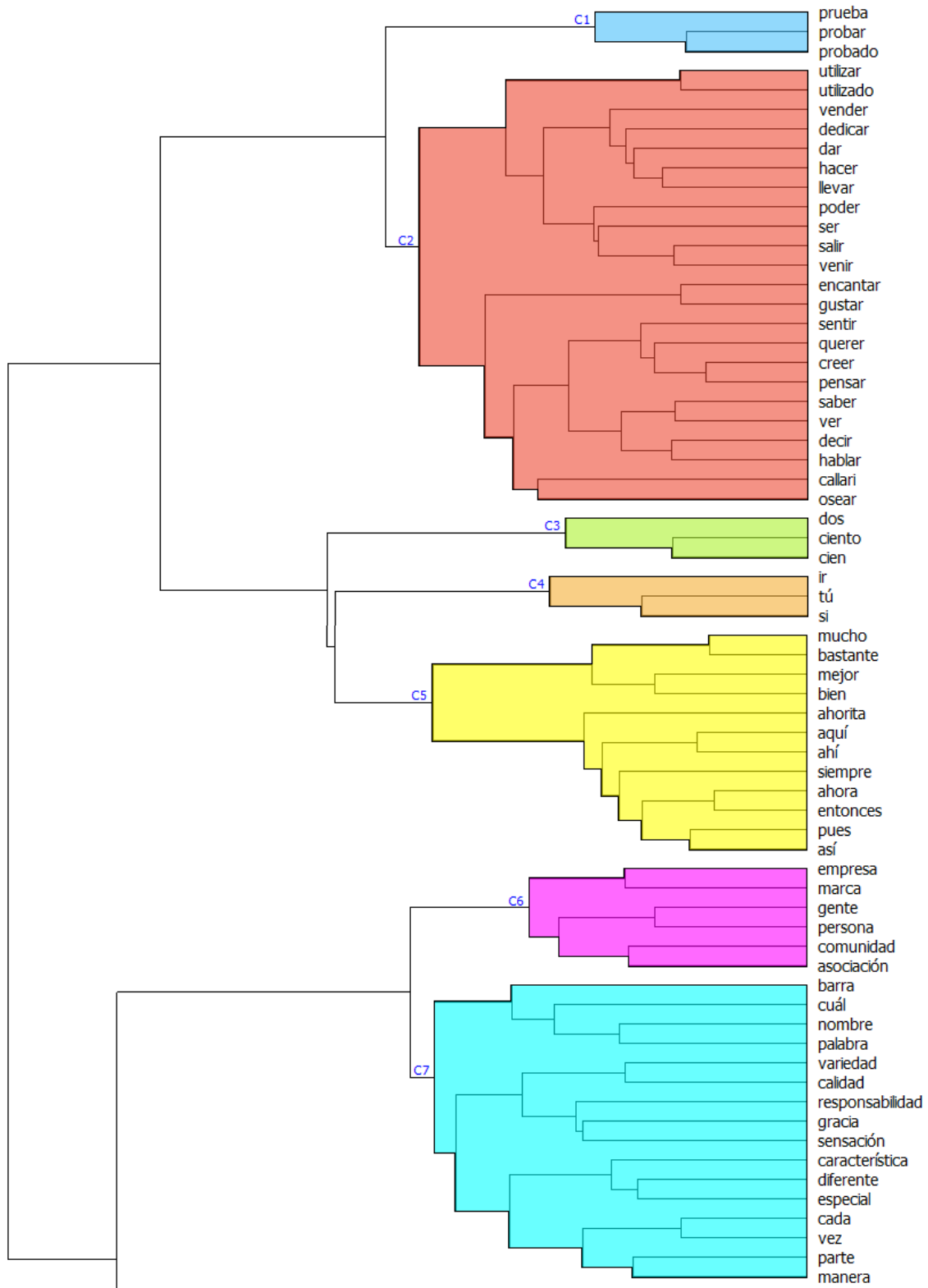
responsable de production : *Miseal* ; et le nom de lieu de production : *Shandia, Talag*. Le concept de *Chakra* tel que nous l'avons défini dans le paragraphe 1.1.2. *L'écologie du langage comme rapport entre langue et environnement naturel*, fait son apparition laissant dans le discours l'empreinte de la spécificité d'un système de culture du cacao, situant ainsi le discours dans un écosystème linguistique particulier.

L'obtention des regroupements hiérarchiques des 100 mots les plus fréquemment utilisés, selon le même processus de vectorisation et de calcul de distance réalisé dans le programme Orange (*Figure 33 : Regroupement hiérarchique des 100 premiers mots du sous-corpus professionnel producteur de chocolat*), offre comme premier résultat deux regroupements qui invitent le destinataire du message à goûter le produit (*prueba, probar, probado*). Cette production discursive s'inscrit dans une stratégie de promotion du produit, une stratégie de vente (*vender, llevar*), en assurant à ceux qui les écoutent qu'ils vont être enchantés (*encantar, gustar, sentir, querrir*). Il apparaît donc, dès les premiers regroupements, un changement de stratégie pour convaincre le destinataire du message : au lieu de la description d'un processus de fabrication durable et organique, des caractéristiques du cacao, de son origine et de ses grandes qualités de *fino de aroma*, le producteur de chocolat invite à goûter son produit. Pour ne pas généraliser et considérer que ce choix discursif est la seule et unique stratégie de promotion qui caractérise le discours du producteur de chocolat en Équateur, il est important de resituer la production langagière dans son écosystème linguistique. En effet, ce choix stratégique n'est sûrement pas étranger au fait que les entrevues ont été réalisées au salon du chocolat devant des stands où les produits étaient exposés pour être dégustés. Les termes du sixième regroupement informent que les personnes sont étroitement liées aussi bien avec l'entreprise qu'avec les communautés. La compréhension de ce regroupement ne peut se faire sans revenir à l'écosystème linguistique. Les producteurs de chocolat, dans leur grande majorité, achètent le cacao aux communautés ou aux associations de communautés. Ils ont donc besoin de garantir aussi bien la qualité du cacao (huitième regroupement : *rico, fino, café, chocolate, cacao, vino, sabor, aroma*) que la réalité d'un processus organique (*orgánico*), tout en négociant le prix d'achat de

la matière première pour assurer la meilleure marge financière et être compétitifs sur le marché du chocolat (dixième regroupement). La tablette (*barra*) de chocolat doit être de qualité (*calidad*), doit offrir une grande variété (*variedad*) pour assurer une sensation caractéristique, différente et spéciale (*sensación característica, diferente et especial*).

Les septième et neuvième regroupements sont directement liés aux questions de l'entrevue, il est logique de retrouver ces termes ainsi regroupés considérant que les entrevues ont été réalisées à partir d'un guide d'entretien validé (*Annexe 4 : Guide d'entretien pour les entrevues en espagnol*). Les adverbes du cinquième regroupement font référence à un changement de situation indiquant qu'aujourd'hui elle s'est améliorée, il y a donc de l'optimisme dans le discours des producteurs de chocolat. Il est important de signaler que les entrevues ont été réalisées avant la crise sanitaire internationale du COVID.

Il est également intéressant d'analyser les regroupements hiérarchiques des mots-clés du sous-corpus professionnel producteur de chocolat. En effet, ces mots-clés sont censés être représentatifs de la terminologie spécialisée du discours considéré selon leur indice de spécificité (Campos *et coll.*, 2020 : 261-271).





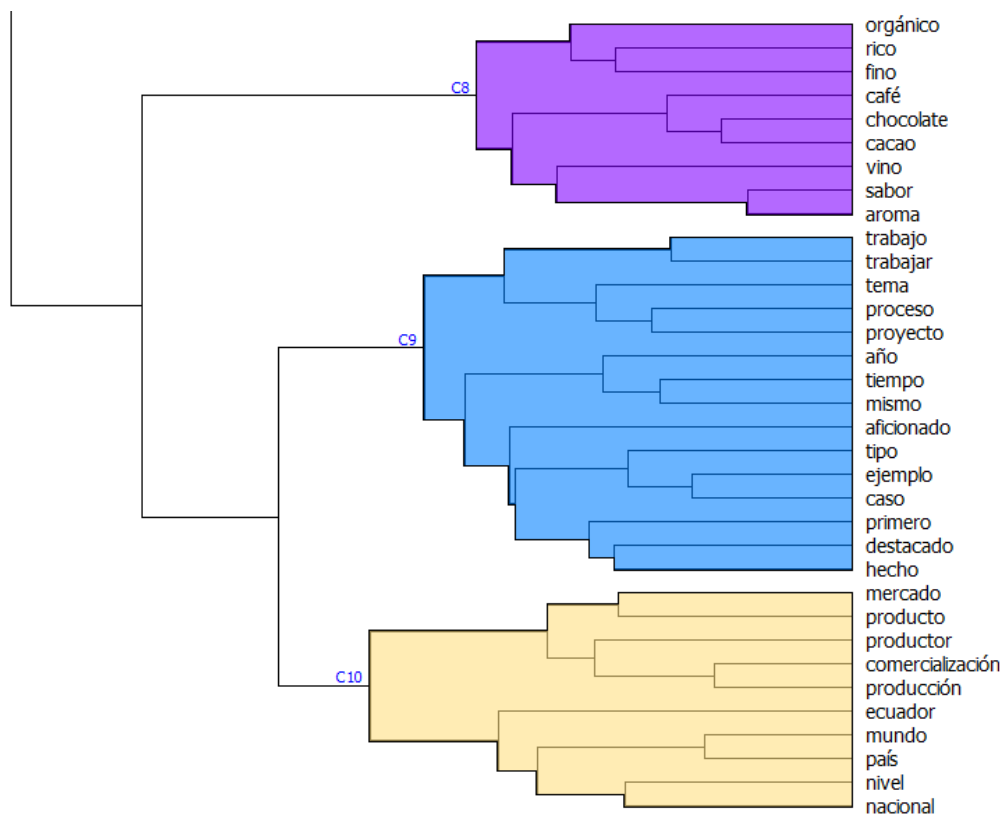
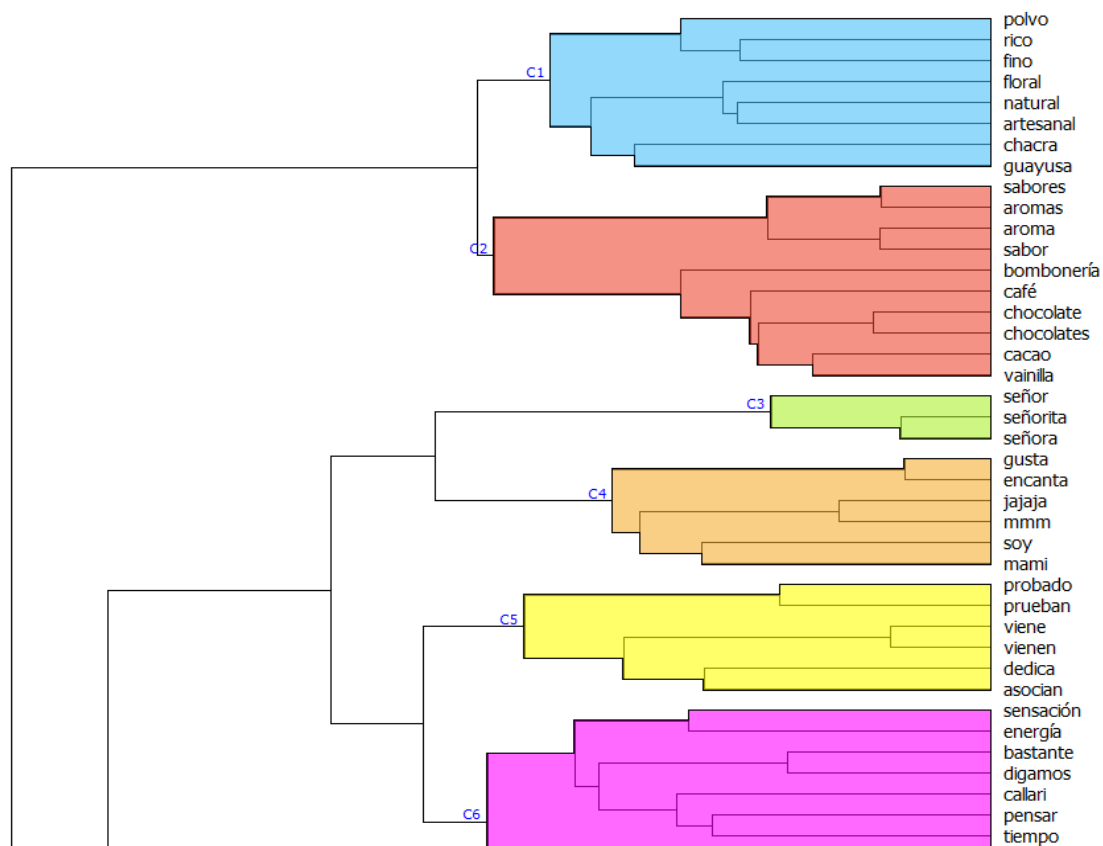


Figure 33 : Regroupement hiérarchique des 100 premiers mots du sous-corpus professionnel producteur de chocolat (Orange)  
 Source : élaboré par nous

Les deux premiers regroupements réunissent les termes qui caractérisent le plus souvent le chocolat. Le mot *polvo* décrit la forme physique. Selon les producteurs de chocolat, *rico* est le mot récurrent que les clients utilisent pour exprimer leurs sensations (*Annexe 12 : Concordances sous-corpus producteur chocolat (TXM)*). Les concordances permettent de vérifier l'utilisation constante du nom qui caractérise le type de cacao équatorien à partir duquel est fabriqué le chocolat : *cacao nacional fino de aroma*. Les termes *café*, *vainilla*, *guayusa* ne sont pas de descripteurs du chocolat, mais des produits qui sont mis sur le marché par certains fabricants de chocolat. Les concordances du terme *artesanal* mettent en évidence le processus de transformation de certains producteurs industrialiser leur production démontrant ainsi que l'industrie de fabrication du chocolat est récente en Équateur alors que l'exportation du cacao a plus de deux cents ans de tradition. La dualité entre *sabor* et *sabores* est intéressante, car

chaque terme a son propre locuteur. En effet, les concordances mettent en évidence que *sabor* est utilisé dans les questions alors que *sabores* est présent dans les réponses des producteurs de chocolat. Une grande majorité des producteurs de chocolat se refusent à réduire leur produit à une seule saveur, utilisant des expressions comme : *un mundo de sabores*, *variedad de sabores*, *bombonería de sabores*. En Espagnol, le mot *bombonería*, comme le mot *bombon*, se réfère à toutes formes de confiserie au chocolat en passant par la truffe au chocolat avec de la liqueur. L'expression *bombonería de sabores* est une analogie entre la variété de confiseries et la variété de saveurs sous-entendant qu'il existe autant de saveurs que de confiserie (au chocolat).



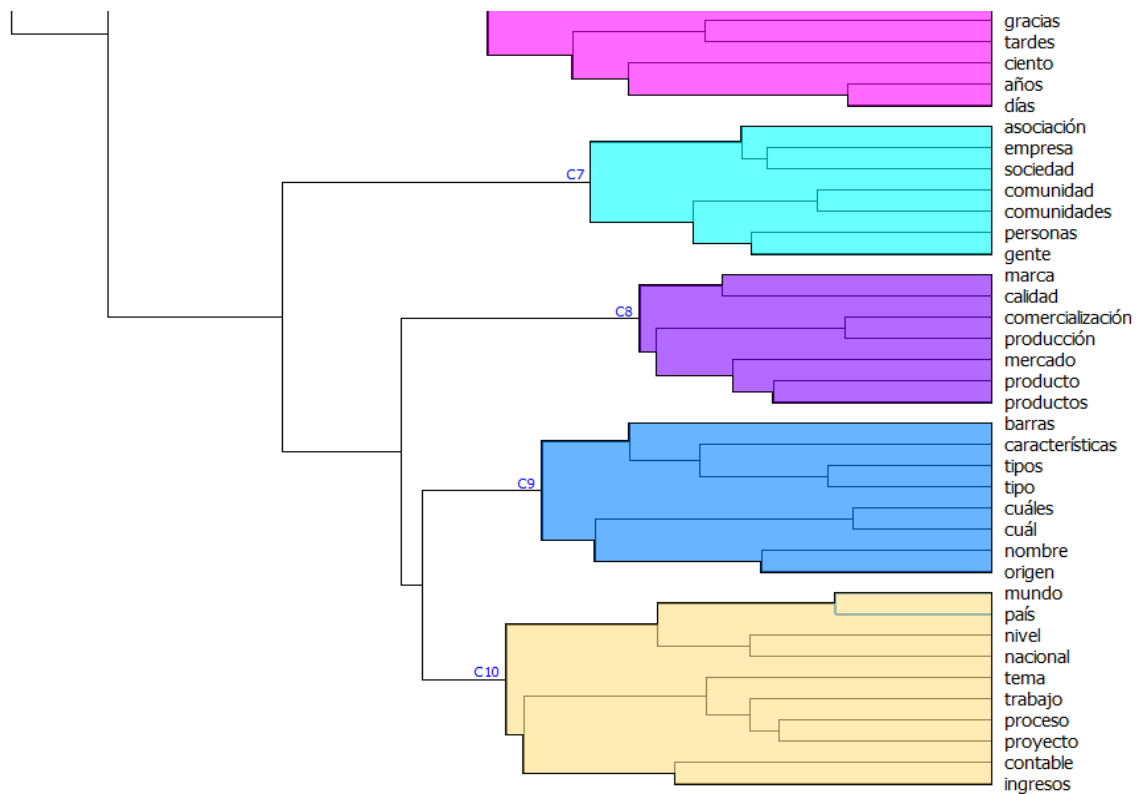


Figure 34 : Regroupement hiérarchique des mots-clés du sous-corpus professionnel producteur de chocolat (Orange)  
 Source : élaboré par nous

Cette précision démontre, une fois de plus, l'importance de l'écosystème linguistique qui suppose que les locuteurs partagent des connaissances aussi bien linguistiques que socioculturelles pour appréhender pleinement le discours partagé.

Il est possible d'observer une dualité comparable entre *aroma* et *aromas* ; le premier faisant partie intégrante du nom du cacao ; le deuxième décrivant le chocolat allant jusqu'à l'hyperbole : *...son aromas de sabores....*

Les troisième, cinquième et neuvième regroupements sont étroitement liés aux questions et aux salutations du début de l'entrevue : *buenas tardes señor, buenos días señorita, buenas tardes señora ¿Cuál es su nombre?, ¿cuál es el nombre de su empresa?*. Le quatrième regroupement répond à la question : *¿Qué dicen las personas que lo prueban?* Plus que des descripteurs sensoriels, la réponse se focalise sur des réactions : *les gusta, les encanta, les llama la atención* (Annexe 12 : Concordances sous-

corpus *producteur chocolat* (TXM)). Le sixième regroupement concerne aussi bien les questions que les réponses. Le producteur de chocolat éprouve des difficultés pour décrire les sensations ressenties. À la question *¿Qué sensación tiene cuando lo prueba?*, les réponses sont hésitantes et peu précises: *una sensación indescriptible, la sensación de que se derrite, primera sensación mi abuelita, un tipo de sensación de felicidad, una sensación de un de un chocolate, esa sensación más de alegría, una sensación de placentero de felicidad de alegría, esa sensación de bueno*, la description reste floue et difficile à mettre en discours. Les septième, huitième et dixième regroupements sont directement en relation avec l'organisation aussi bien de l'entreprise pour la production que du marché pour la commercialisation (*Annexe 12 : Concordances sous-corpus producteur chocolat* (TXM)) : *nosotros trabajamos con veintinueve comunidades en la Amazonía específicamente en la zona de Archidona, la sociedad por afuera conozca como como la comunidad produce este tipo de producto sano, manera nos permite como le digo localmente abrir mercados y también exportar a través de nuestros socios comerciales, hacia la gente y de poder entregar al mercado un producto tan exclusivo es una sensación gratificante, mi caso yo manejo el área comercial tanto nacional como internacional, combinaciones son muy atractivas para el el mercado nacional y en el mercado del exterior en cambio ellos buscan altos contenidos en cacao.*

Dans la représentation graphique de la projection de l'échelonnement multidimensionnel de mots appartenant au sous-corpus professionnel du chocolat dont la fréquence est supérieure à dix (*Figure 35 : Agrandissement de la projection bidimensionnelle (MDS) des mots du sous-corpus producteur de chocolat dont la fréquence est supérieure à dix (Orange)*), les septième, huitième et neuvième regroupements se trouvent être le plus souvent en interaction reliant les questions de l'entrevue (septième et neuvième regroupements) directement à la qualité du cacao (huitième regroupement).





comme les substantifs qui décrivent les produits obtenus pendant le processus de transformation (*plantas, pasta, mucilago, manteca, nibs, polvo*).

<b>Noms</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	<b>Adjectifs</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
aroma	55	39	17	3,7	fino	44	34	17	1,4
cacao	314	103	11	5,8	nacional	26	16	6	2,1
plantas	14	9	4	3,2	destacado	6	5	3	4,6
maceración	4	4	3	5,3	<b>Verbes</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
origen	14	8	3	3,4	producir	3	3	2	2,7
variedad	12	7	3	3,7	utilizar	3	3	2	3,3
productores	32	13	3	4,0	adquiere	3	3	2	4,0
producción	33	13	3	2,8	investigar	3	3	2	6,7
pasta	7	5	3	4,6	encuentran	3	3	2	5,7
mucilago	7	5	3	2,4	generamos	3	3	2	4,0
manteca	7	5	3	1,8	tiene	179	45	2	5,0
nibs	5	4	2	1,0	dedica	22	9	2	5,8
perfiles	5	4	2	6,8	<b>Noms Propres</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
variedades	14	7	2	3,3	Ríos	3	3	2	4,3
recetas	3	3	2	6,3					
polvo	11	6	2	4,7					

Tableau 13 : cooccurrence de cacao dans le sous-corpus professionnel producteur de chocolat (TXM)

Source : élaboré par nous

<b>Noms</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	<b>Verbes</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
comercialización	46	35	15	2	dedican	17	14	7	6
tiempo	38	18	4	7	probado	27	18	7	2
chocolate	353	99	4	6	dedicaron	15	12	6	5,5
salón	5	5	3	1,2	hace	88	33	4	6
licor	5	5	3	2,2	comer	4	4	3	3
prueba	38	17	3	3	gusta	48	19	3	5
sensaciones	9	6	2	6	pensar	23	11	3	5

leche	15	8	2	2,5	dedicó	5	4	2	4
sensación	52	19	2	6	esperar	3	3	2	5
<b>Adjectifs</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	tiene	179	50	2	4,9
mejor	22	10	2	3	<b>Noms Propres</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
oscuro	3	3	2	7	Callari	30	13	2	6

Tableau 14 : cooccurrence de chocolate dans le sous-corpus professionnel producteur de chocolat (TXM)

Source : élaboré par nous

Dans le tableau des cooccurrents de *chocolate* les résultats sont directement liés aux questions de l’entrevue (termes surlignés). Le substantif *salón* est en relation avec le lieu des entrevues (salon du chocolat) ; *licor* (ou *pasta*) est un produit du processus de fabrication du chocolat et *leche* un ingrédient. *Mejor* et *oscuro* sont les deux seuls adjectifs qui présentent une cooccurrence notable avec *chocolate* sans pour autant être représentatifs d’un discours sensoriel.

Le programme Orange permet l’obtention de la représentation graphique du réseau sémantique des termes dont la fréquence est supérieure à 35, dans une fenêtre de deux mots à gauche et à droite, des cooccurrences dont la fréquence est supérieure à 20 (*Figure 37* : Réseau sémantique des principales cooccurrences du sous-corpus professionnel producteur de chocolat (Orange)). Les termes utilisés dans les questions de l’entrevue sont surlignés en jaune. Les réponses restent orientées autour de la commercialisation du produit dont les qualités sont intrinsèquement liées au cacao *fino de aroma* en harmonie avec le discours du sous-corpus ambiant.



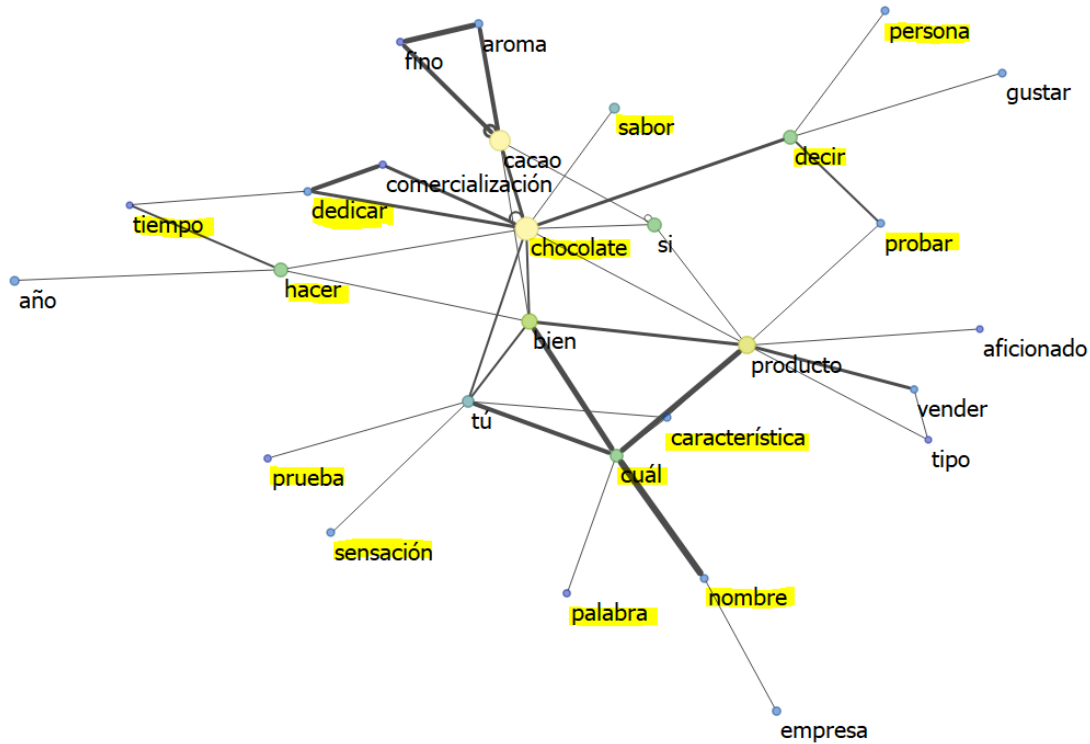


Figure 37 : Réseau sémantique des principales cooccurrences du sous-corpus professionnel producteur de chocolat (Orange)  
 Source : élaboré par nous

Bien que le sous-corpus professionnel producteur de chocolat présente de nombreuses similitudes avec le sous-corpus ambiant, l'analyse des sentiments (Sentiment Analysis) du programme Orange présente une moyenne de quatre sur l'échelle des sentiments et 50% des textes se situent entre trois et cinq et demi (Figure 38 : Représentation de l'analyse de sentiment des textes du sous-corpus professionnel producteur de chocolat (Orange)). Ce résultat met en évidence que les professionnels ont un discours deux fois plus positif que le discours ambiant publié sur Internet quand ils parlent de leurs produits chocolatés. Encore une fois, ce résultat n'est pas surprenant, car, quel que soit l'objectif de la communication, leur intention de communication reste la promotion et la vente de leurs produits.

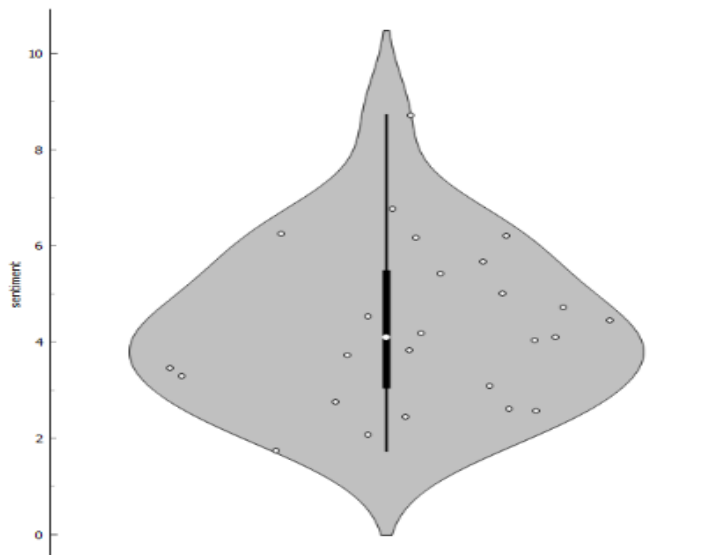


Figure 38 : Représentation de l'analyse de sentiment des textes du sous-corpus professionnel producteur de chocolat (Orange)  
Source : élaboré par nous

### 3.2.2. Moule discursif représentatif du discours professionnel du producteur de cacao en espagnol

L'écosystème linguistique dans lequel a été produit le discours des producteurs de cacao est différent de celui d'un salon du chocolat ou d'une entreprise ; c'est au pied des cacaotiers, dans les fermes de production que ces entretiens ont été réalisés en suivant le guide d'entretien correspondant (*annexe 4*). L'amplitude du sous-corpus professionnel des producteurs de cacao est comparable à celle du sous-corpus professionnel des producteurs de chocolat avec une densité lexicale légèrement inférieure (*Tableau 15 : Paramètres du sous-corpus professionnel producteur de cacao*). Les locuteurs sont, dans leur grande majorité, des personnes qui ont un accès difficile à l'éducation, dont la langue maternelle est très souvent le kichwa. Cependant cette réalité se reflète plus dans la structure de l'espagnol produit que dans le chiffre de densité lexicale. Par exemple, dans la contribution C-012, le terme *vía* est ainsi prononcé alors que le terme correct est *veía* ; de même pour le terme *mimo* qui devrait être prononcé *mismo*. Il existe aussi une variation dans la morphologie du vocabulaire utilisé. En effet,

des termes hispanisés empruntés au kichwa apparaissent dans le discours comme *guambra* issu de *wanra* qui désigne un gamin ou une gamine, ainsi que l'utilisation fréquente des suffixes –ito/a (*poquito, pistolita, guambrito, solito*) et –azo/a (*salivazo*) ce qui a un impact direct sur le calcul de la densité lexicale sans pour autant augmenter la variété lexicale du discours.

*C-011 : entrevistado: entonces eso le traje un poquito no más y con ese de ahí una pistolita en la bomba ahí le deje*

*C-012 : entrevistado: vía guambrito mimo vía entonces digo yo si he de poder hacer yo puedo hacer así que un día solito estaba haciendo le taste y si no <...> faltado poquito más aplastado*

*C-013 : entrevistado: y se llena el salivazo esa una un bicho que es baboso*

Il est aussi important de considérer que dans le calcul de la densité lexicale sont inclus les termes utilisés pour poser les questions, ce qui uniformise le résultat entre les deux sous-corpus.

sous-corpus SCP cacao espagnol	Token	Vocabulaire	Densité lexicale sur 1000 tokens aléatoires
Nombre total	28683	2172	0.328

*Tableau 15 : Paramètres du sous-corpus professionnel producteur de cacao*  
*Source : élaboré par nous*

L'écart observé entre le sous-corpus ambiant et le sous-corpus professionnel de producteur de cacao est aussi expliqué par les conditions de production du discours entre production asynchrone et synchrone. Pour les mêmes raisons d'extension du sous-corpus professionnel producteur de cacao, au lieu des cent premiers lemmes de chaque partie de discours considérée, le *Tableau 16 : Lemmes du sous-corpus professionnel producteur de cacao* dont la fréquence est supérieure à 10 réunit les lemmes dont la fréquence est supérieure à dix tout en observant une classification selon

les parties du discours. Au niveau des substantifs, aussi bien dans le *Tableau 11* que dans le *Tableau 16*, cacao et chocolat semblent indissociables, que le thème soit l'un ou l'autre, l'un implique l'autre et vice-versa.

<b>[espos="NC"]: eslemma / F</b>							
entrevistador	602	sensación	56	agricultor	19	libra	12
cacao	450	plátano	50	cosa	19	medio	12
chocolate	407	día	49	café	18	participación	12
sabor	102	persona	49	planta	16	cambio	11
año	100	tipo	44	señora	16	compañero	11
producción	98	agricultura	38	hectárea	15	dinero	11
trabajo	96	fecha	37	precio	15	finca	11
producto	88	ejemplo	36	casa	14	molino	11
comunidad	87	aroma	32	fábrica	14	asociación	10
tiempo	71	familia	31	pepa	14	grupo	10
nombre	63	gracia	31	campo	13	ok	10
responsabilidad	63	maíz	27	taller	13	yuca	10
característica	62	gente	26	zona	13		
vez	60	señor	23	dólar	12		
palabra	57	hijo	21	dulce	12		
<b>[espos="ADJ"]: eslemma / F</b>							
bueno	269	solo	26	dulce	17	sabroso	14
mismo	59	fino	21	mucho	17	grande	12
rico	57	natural	21	bueno mejor	16	limpio	11
claro	47	nuestro	20	perfecto	15	orgánico	10
<b>[espos="V.*"]: eslemma / F</b>							
entrevistar	1003	dar	67	seguir	32	crear creer creer	15
ser	652	utilizar	63	saber	29	cultivar	15
hacer	381	trabajar	59	agradecer	28	dejar	15
tener	306	hablar	58	producir	26	partir	14
haber	305	elaborar	56	gustar	25	llevar	13
decir	221	venir	56	salir	25	pasar	13

dedicar	190	sembrar	48	comprar	23	salar	13
estar	138	poner	43	mantener	21	vivir	13
probar	118	vender	41	preparar	20	comenzar	12
vendar vender	110	ver	41	llamar	19	ganar	12
poder	109	ir	38	secar	19	tomar	12
pensar	84	definir	37	querer	17	limpiar	11
osear	77	sacar	32	aprender	15	entrever	10
<b>[espos="ADV"]: eslemma / F</b>							
ya	422	como	86	mucho	32	siempre	14
más	276	muy	79	menos	20	acá	11
bien	179	también	79	nunca	20	después	11
entonces	117	aquí	66	cómo	19	luego	11
así	99	ahora	37	poco	19	cuándo	10
ahí	93	casi	32	bastante	18	todavía	10

Tableau 16 : Lemmes du sous-corpus professionnel producteur de cacao dont la fréquence est supérieure à 10 (TXM)

Source : élaboré par nous

De *sabor* à *sensación*, se sont les treize noms utilisés dans les questions qui sont représentés. En *annexe 13*, il est possible de vérifier que *sabor* est aussi bien utilisé par celui qui pose les questions que par celui qui y répond, il n'existe pas de différence faite par les producteurs de chocolat entre *sabor* et *sabores*. Dans le discours du producteur de cacao, les cooccurrents de *sabor* son *natural*, *rico*, *distinto*, *agradable*, *elegante*, *dulce*. Il est surprenant de constater que le naturel et l'élégance aient un goût particulier. Il est possible d'imaginer que le locuteur souhaite que les représentations sociales qui pour lui ont une forte valeur ajoutée, comme le naturel et l'élégance, soient partie intégrante de son cacao à tel point qu'elles laissent un goût particulier au chocolat. En approfondissant l'analyse, dans les concordances, le terme *sabor* est aussi associé à : *increíble*, *aromático*, *excelente*, *delicioso*, *amargo*, *fino*, *especial*, qui restent des adjectifs communs pour qualifier une impression générale sans vraiment être des descriptifs sensoriels spécifiques du cacao ou du chocolat. Dans le tableau des

concordances de *sabor* (*annexe 13; Error! No se encuentra el origen de la referencia.*), certaines expressions offrent une compréhension plus étendue du contexte écolinguistique des producteurs de cacao : *como que nunca antes hemos probado, no sé no me acuerdo, no se como puedo decir eso, sabor de cacao, bien fragante como a chocolate, de que uno se puede tener ánimo, de que uno se siente más animoso, sabor de chocolate*. Certains producteurs de cacao n'ont jamais goûté le chocolat fabriqué à partir de leur cacao (voir concordance de *nunca* en *annexe 13; Error! No se encuentra el origen de la referencia.*). Ils l'appellent *pepa de oro* alors qu'ils ne connaissent ni la couleur, ni l'odeur, ni la saveur de cet or. Ils éprouvent de grandes difficultés pour décrire les sensations que procure un chocolat exporté, car il ne fait pas partie de leur réalité, ce qui explique la pauvreté des adjectifs utilisés. Le chocolat qu'ils consomment sous forme de boisson préparée de manière artisanale n'est pas comparable au produit exporté à l'international. Comme le confirment les autres substantifs les plus fréquemment utilisés, le producteur de cacao est plus un agriculteur qui cultive différents produits comme le *plátano, maíz, café, yuca* dans sa *finca* qui le plus souvent est familiale qu'un producteur de cacao exclusif qui vendrait le chocolat de sa ferme en s'impliquant dans tout le processus de fabrication. Les verbes les plus fréquents confirment cette analyse en décrivant les processus pour obtenir une fève prête pour être vendue au producteur de chocolat : *utilizar, trabajar, elaborar, sembrar, sacar, producir, mantener, preparar, secar, cultivar...*

En ce qui concerne les adverbes, les plus fréquents sont des marqueurs de l'oralité : *ya, bien, así, ahí* ; tout comme *osea* qui, bien qu'annoté comme le verbe *osear*<sup>50</sup>, est la traduction de « c'est-à-dire » (*o sea*). D'autres marqueurs de l'oralité apparaissent dans le discours comme : *m, aa, eh, ah, jajaja, jejeje* qui selon les conventions de transcription peuvent exprimer la confirmation, l'hésitation, ou le rire (*Figure 39*).

---

<sup>50</sup> Le verbe *osear* est de moins en moins utilisé, il signifie faire peur, faire fuir.



transport des produits de la chakra à la route), *papa china (colocasia esculenta)*, *naranjilla (solanum quitoense)*.

Word	Yake	Word	Yake	Word	Yake	Word	Yake
fecha	0,18	chakra	0,02	medio	0,01	necesitamos	0,01
entrevista	0,17	alvarado	0,02	soy	0,01	caiza	0,01
nombre	0,16	viene	0,02	hacemos	0,01	cerda	0,01
cuál	0,16	andi	0,02	fino	0,01	podas	0,01
dedica	0,15	rico	0,02	probado	0,01	hijos	0,01
producción	0,11	palabras	0,02	zona	0,01	comunidades	0,01
días	0,11	maíz	0,02	maría	0,01	comercial	0,01
comunidad	0,10	responsabilidades	0,02	mis	0,01	responsabilidad	0,01
trabajo	0,09	pensar	0,02	shiguarcocha	0,01	perfecto	0,01
señor	0,07	características	0,02	aroma	0,01	caballos	0,01
señora	0,07	dedicó	0,02	agricultor	0,01	tapuy	0,01
sabor	0,06	jejeje	0,02	vendo	0,01	grupo	0,01
cacao	0,06	dedico	0,02	tipo	0,01	siembra	0,01
chocolate	0,06	rey	0,02	entrevistador	0,01	productos	0,01
años	0,05	osea	0,02	andy	0,01	experiencia	0,01
producto	0,05	dulce	0,02	asociación	0,01	niños	0,01
plátano	0,04	ahorita	0,01	chocolates	0,01	china	0,01
agricultura	0,04	encuesta	0,01	vargas	0,01	papa	0,01
vende	0,04	tanguila	0,01	pepa	0,01	shiguango	0,01
cuáles	0,04	docente	0,01	cincuenta	0,01	naranjilla	0,01
trabajar	0,03	tardes	0,01	elaborado	0,01	controlar	0,01
yuca	0,03	hablar	0,01	siembran	0,01	esperanza	0,01
tiempo	0,02	español	0,01	fanny	0,01	taller	0,01
natural	0,02	claro	0,01	calapucha	0,01	campo	0,01
probó	0,02	fábrica	0,01	aguinda	0,01	persona	0,01

Tableau 17 : 100 premiers mots-clés du sous-corpus professionnel producteur de cacao en espagnol (Orange)

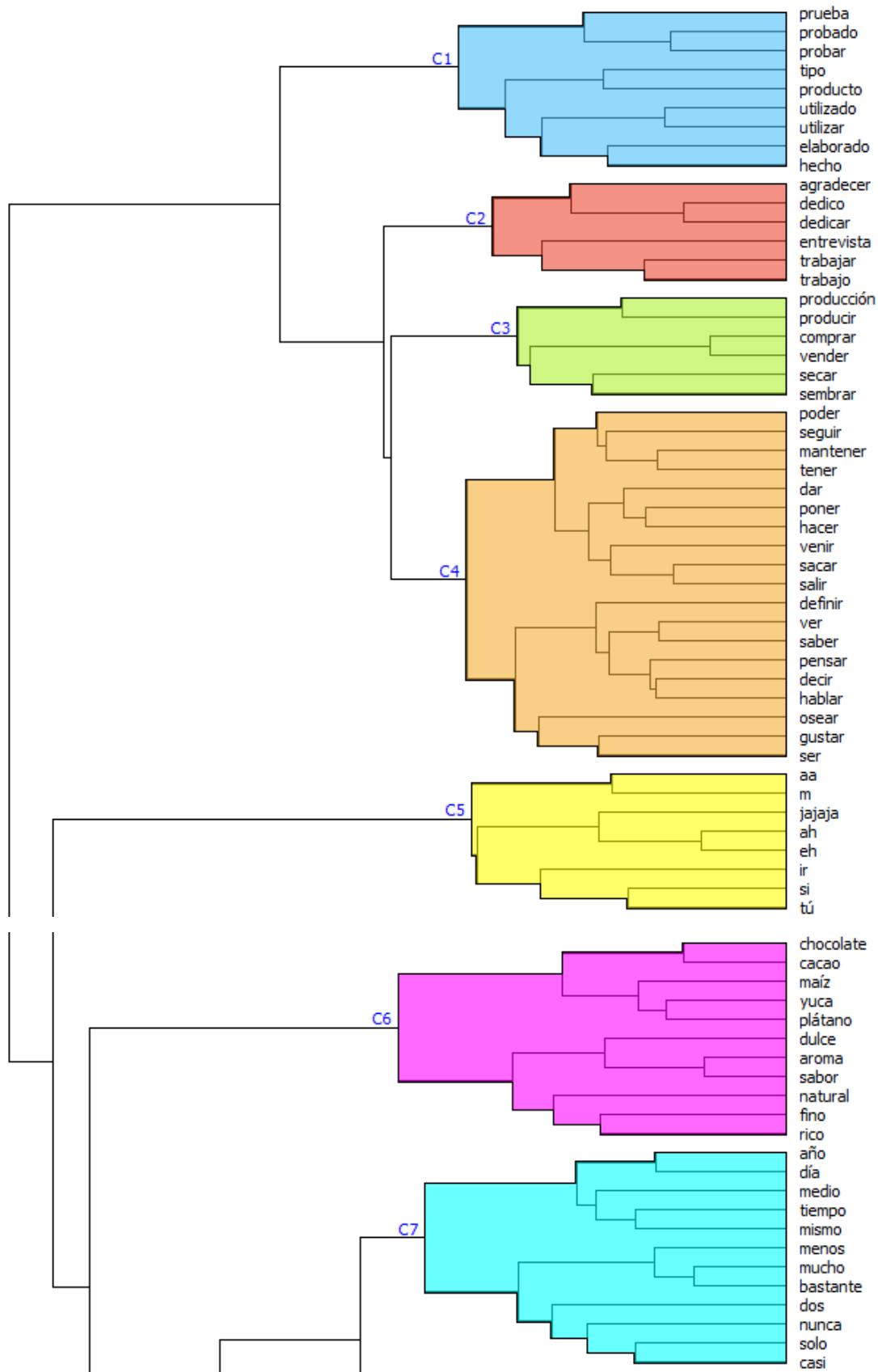
Source : élaboré par nous

Une des participantes aux entrevues utilise le terme *docente* pour souligner le fait que sa profession est maîtresse d'école à l'école primaire de sa communauté et quand elle finit son travail d'enseignante, elle s'occupe de ses plantations.



En considérant la similitude des deux listes de mots, seul le regroupement hiérarchique des cent premiers mots sera partagé, car il a le mérite de ne pas présenter un regroupement autour des noms propres (*Figure 40* : Regroupement hiérarchique des 100 premiers mots du sous-corpus professionnel producteur de cacao (Orange)). Les deux premiers regroupements (C1 et C2) sont directement liés aux questions posées lors de l'entrevue. Les deux suivants (C3 et C4) sont directement liés aux réponses, le regroupement C3 décrit les étapes de production du cacao : *sembrar, producir, secar, comprar / vender*. Le regroupement C4 est plus en relation avec la raison pour laquelle ils cultivent le cacao : *poder, seguir, mantener, tener*. La culture du cacao apparaît comme un moyen de subsistance au même titre que les autres produits qu'ils cultivent comme la yuca (manioc d'Amérique de sud), le maïs, les bananes plantains, qui se retrouvent dans le sixième regroupement auxquels il est possible d'ajouter les produits identifiés dans les mots-clés : *papa china, naranjilla*. Les autres termes du quatrième regroupement font partie de la question. Le cinquième regroupement réunit les onomatopées : *aa, m, jajaja, ah, eh* ; le lemme *ir* pour le terme *vaya* qui s'utilise comme un marqueur de l'oralité. Le terme *si* pour répondre aux questions ; et le déictique *tú* confirmant ainsi que le cinquième regroupement réunit les marqueurs de l'oralité du discours étudié. Le sixième regroupement, en plus de mentionner les autres produits principaux cultivés dans la *chakra*, réunit le peu de termes qui peuvent être associés au discours sensoriel : *dulce, aroma, sabor, fino, rico* tout en conservant la référence à l'expression *fino de aroma*. Les septième et huitième regroupements, pour présenter en majorité des adverbes, des adjectifs interrogatifs ou cardinaux, n'apportent pas une information significative. L'association de *sensación* et *característica* dans le neuvième regroupement correspond aux questions de l'entrevue. Le dixième regroupement confirme que l'écosystème linguistique du producteur de cacao s'inscrit dans un contexte communautaire où l'agriculture est une entreprise familiale qui permet la production de différents produits parmi lesquels se trouve le cacao. Les cultivateurs de cacao s'emploient à fournir une fève répondant aux standards de qualité du producteur de chocolat sans vraiment connaître le produit final fabriqué à partir de leur cacao. Le

chocolat n'étant pas la priorité, il est peu surprenant que le discours sensoriel soit très limité dans le discours de producteur de cacao.



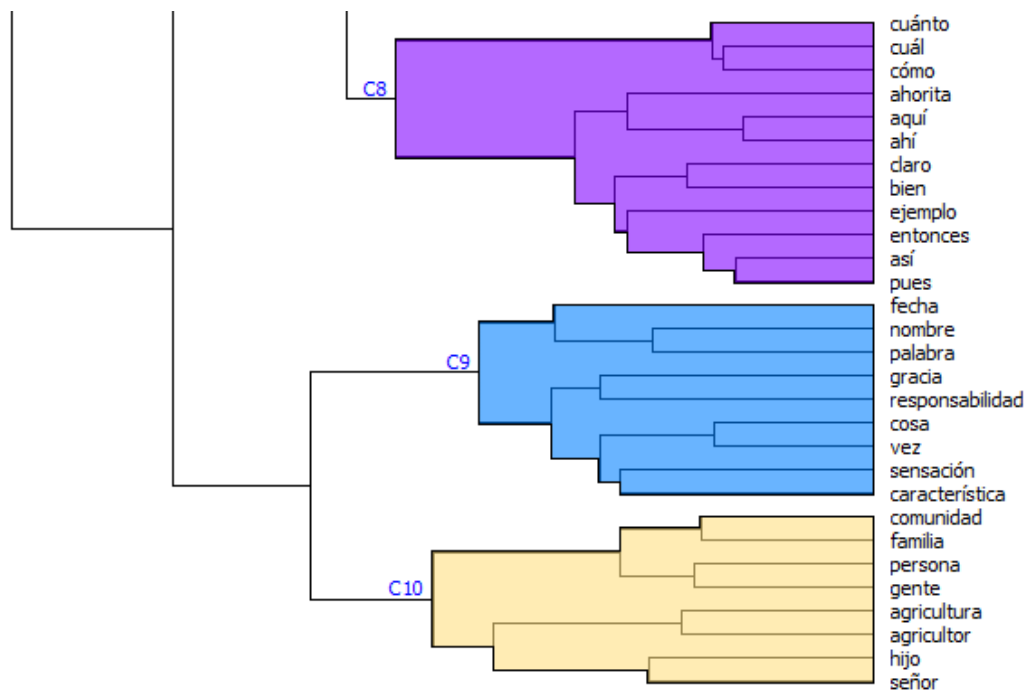


Figure 40 : Regroupement hiérarchique des 100 premiers mots du sous-corpus professionnel producteur de cacao (Orange)  
Source : élaboré par nous

Dans la représentation graphique de la projection de l'échelonnement multidimensionnel des cent premiers mots appartenant au sous-corpus professionnel du producteur de cacao (Figure 41), il n'est pas surprenant que les regroupements représentatifs des questions (C1, C2 et C9) soient en relation directe avec ceux qui sont représentatifs des réponses (C3, C4 et C6) confirmant ainsi la structure même du discours étudié. Cependant, la présence de cinq termes du dixième regroupement (*comunidad, familia, agricultor, agricultura, persona*) au centre des interactions est une référence directe à la réalité dans laquelle le locuteur est immergé, se définissant lui-même comme un agriculteur communautaire qui vend du cacao comme il vend les autres produits de son agriculture. C'est aussi une des raisons pour lesquelles son discours est centré sur le processus d'obtention de la fève de cacao, qu'il connaît, et non sur les sensations produites par le chocolat fabriqué à partir de sa *pepa de oro*.

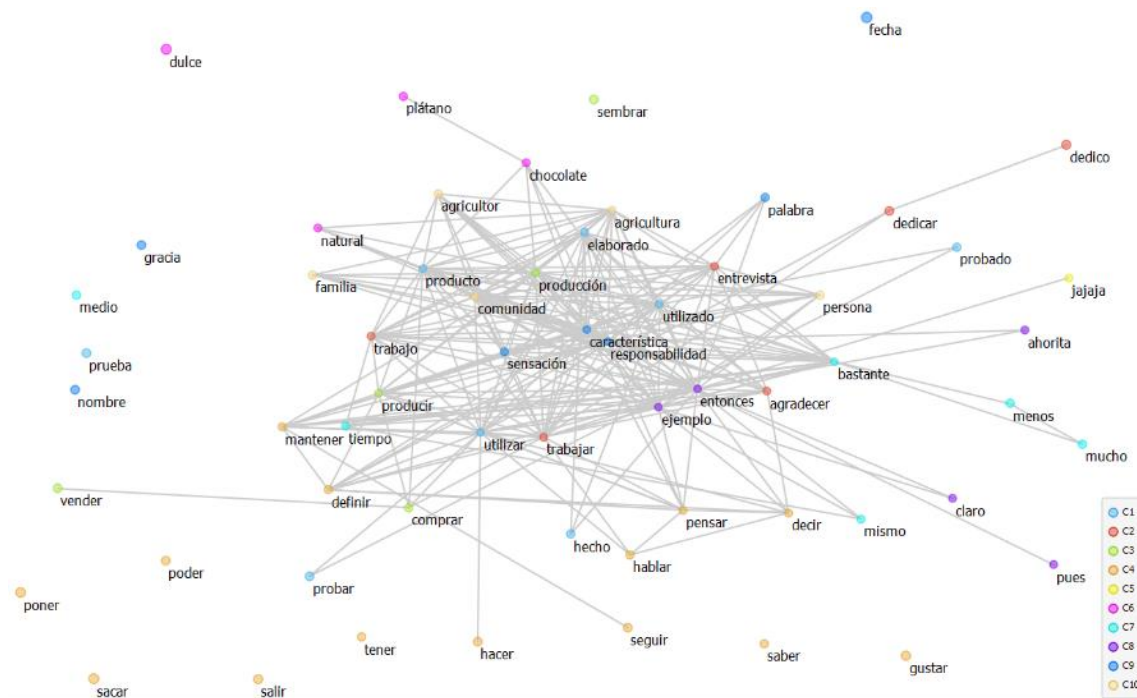


Figure 41 : Agrandissement de la projection bidimensionnelle (MDS) des cent premiers mots du sous-corpus producteur de chocolat (Orange)  
 Source : élaboré par nous

En observant les cooccurrences des termes : cacao (Tableau 18 Tableau 18 : cooccurrence de cacao dans le sous-corpus professionnel producteur de cacao Tableau 13 : cooccurrence de cacao dans le sous-corpus professionnel producteur de chocolat), il est possible de constater la présence de l'adjectif *fino* dont les concordances (Tableau 19 Tableau 18 : cooccurrence de cacao dans le sous-corpus professionnel producteur de cacao Tableau 13 : cooccurrence de cacao dans le sous-corpus professionnel producteur de chocolat) confirment la référence à l'expression figée *fino de aroma* qui identifie le type de cacao cultivé et qui à elle seule suffit pour déterminer les arômes et les saveurs : *era un poco más aromático por porque es fino de aroma no, estaba el sabor o sea fino como fino de aroma netamente a cacao.*

Noms	Freq	CoFreq	Score	MDist	Verbes	Freq	CoFreq	Score	MDist
producción	98	87	38	1.8	dedicó	41	40	21	4.9
características	61	43	12	6.2	elaborado	47	39	15	2.9

tiempo	69	45	10	7.0	definir	37	33	14	7.1
plátano	46	32	9	3.3	puede	50	34	9	7.7
maíz	26	19	6	3.7	dedica	98	50	6	5
yuca	31	21	5	3.7	partir	12	11	5	2
cacao	448	147	3	5.4	hace	142	62	5	7
temporada	5	5	2	3.4	vende	66	33	4	3.9
siembra	11	8	2	3.0	sembrar	20	14	4	2.9
años	92	36	2	6.9	dedico	21	14	3	5.2
caña	6	5	2	3.6	tengo	33	19	3	3.8
<b>Adjectifs</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	vendo	17	10	2	5.4
fino	20	13	3	2.6	produce	4	4	2	1.7

Tableau 18 : cooccurrence de cacao dans le sous-corpus professionnel producteur de cacao (TXM)

Source : élaboré par nous

LeftContext	Keyword	RightContext
entiende que es cacao nosotros tenemos	fino	aroma entrevistador: ya entrevistado: el cacao natural es
: ya entrevistado: el cacao natural es	fino	aroma entrevistador: < ah > ya pero características tiene el
elaborado con este choco este cacao de	fino	aroma entrevistado: si ya eso da tiene un sabor bien fragante
se comienza a dañarse en cambio el	fino	aroma cuando tú haces tostar chocolate eso te dura
entrevistado: no el otro entrevistador: el	fino	aroma dura entrevistado: dura eso no tiene nada en cambio
como dos hectáreas y medio es el cacao	fino	fino de aroma y luego hice punto C50 entonces tengo dos
dos hectáreas y medio es el cacao fino	fino	de aroma y luego hice punto C50 entonces tengo dos
señores del MAP una mesa de de cacao	fino	de aroma un evento donde que se hizo < a > las
es puro cacao como tenemos la variedad	fino	de aroma entonces cacao olor a aroma la naturaleza que
me hace pensar que es un producto muy	fino	muy bueno que nosotros lo podemos consumir a nuestro.
semillero y plantamos nosotros plantas de	fino	de aroma del nacional entrevistador: bien entrevistado:
en esas vino dos variedades la una de	fino	de aroma y la otra del CCN 51 entrevistador: bien y
entrevistado: mi cacao es cacao nacional	fino	de aroma entrevistador: ah ya entrevistado: y tengo dos
era un poco más aromático por porque es	fino	de aroma no < jajaja > entrevistador: y ¿ a qué
dos tipos de chocolate chocolate lo que es	fino	de aroma también de trinitario hay una diferencia creo que e
tenía cinco hectáreas de cacao en la	fino	de aroma de aquí de la zona mismo de cacao nacional que
: < eh > estaba el sabor osea	fino	como fino de aroma netamente a cacao entrevistador:
eh > estaba el sabor osea fino como	fino	de aroma netamente a cacao entrevistador: netamente
como primeros productores de cacao de	fino	de aroma entrevistador: ya entrevistado: entonces ellos son l

una buena acogida como como cacao	fino	de aroma entrevistador: claro entrevistado: entonces de esta
-----------------------------------	------	--

Tableau 19 : concordances de *fino* dans le sous-corpus professionnel producteur de cacao (TXM)

Source : élaboré par nous

Les substantifs et les verbes, cooccurrents de *cacao*, confirment le résultat de l'analyse précédente décrivant un discours d'agriculteur dont l'un de ses produits cultivés est le cacao. Le locuteur réduit sa production discursive au processus de culture. La présence de *elaborado* est due à la question posée : *Puede definir las características del chocolate elaborado a partir de su cacao?* Dans la grande majorité des cas, l'agriculteur ne peut pas répondre à la question (Tableau 20), car la fabrication du chocolat ne fait pas partie de son quotidien.

Keyword	RightContext
elaborado a partir de su cacao?	entrevistado: < eh > bueno como como se puede decir bueno yo últimamente como osea no he hecho con mi con mi cacao
elaborado a partir de su cacao?	entrevistado: < m > bueno como le digo no he hecho yo digamos una < ... > un procesamiento legalmente no no no he tratado de hacer
elaborado a partir de su cacao?	entrevistado: si < eh > si lo que pasa que yo me he dedicado como digo es el la la idea la idea y el conocimiento que
elaborado a partir de su cacao?	entrevistado: bueno < m > una vez que fuimos con el con el producto que fuimos llevando en las cabañas Anzu hicimos el el el chocolate con <
elaborado a partir de su cacao?	entrevistado: bueno eso no hemos < hemos > dedicado justamente unos dos talleres que nos dieron los del ministerio mismo osea en grupo hemos tratado de < elaborar
elaborado a partir de su cacao?	entrevistado: no entendí cómo cómo entrevistador: < m > ¿ qué características tiene el chocolate que usted ha hecho con su cacao? en casa por ejemplo entrevistado
elaborado a partir de su cacao?	entrevistado: no se eso si casi yo como persona no me he dedicado hacer los chocolates pienso que debe ser uno de los mejores < jajaja > entrevistador
elaborado a partir de su cacao?	entrevistado: bueno yo al no puedo mentir porque yo no he participado pero en la en esta comunidad si tenemos un grupo de de compañeras que somos casi
elaborado a partir de su cacao?	entrevistado: pero mi cacao no he hecho todavía entrevistador: no he hecho entrevistado: no he hecho entrevistador: pero ha hecho de otro entrevistado: <
elaborado a partir de su cacao?	entrevistado: si el cacao nosotros < sac > cogemos el cacao le secamos al sol así como se como es debido el tratamiento le ponemos < eh >
elaborado a partir de su cacao?	entrevistado: < eeh > chocolate cacao de chocolate esta rico para mi < aa > entrevistador: está rico para usted entrevistado: está rico si entrevistador: y

Tableau 20 : concordances de « elaborado a partir de su cacao? » dans le sous-corpus professionnel producteur de cacao (TXM)

Source : élaboré par nous

En ce qui concerne les cooccurrences du terme *chocolate* (Tableau 21/ Tableau 14 : cooccurrence de chocolate dans le sous-corpus professionnel producteur de chocolat) obtenues avec le programme TXM, les substantifs et les verbes sont les termes utilisés dans les questions sur les caractéristiques et les impressions que laisse le chocolat élaboré à partir du cacao produit. Ces cooccurrences ne sont pas représentatives du discours du producteur de cacao.

<b>Noms</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	<b>Verbes</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>
características	61	56	28	2.0	puede	50	33	9	4.3
palabras	57	50	23	5.8	tiene	154	71	9	4.0
personas	44	42	23	4.8	tuvo	17	16	8	4.1
sensación	55	46	20	5.4	prueba	23	19	8	1.5
sabor	100	62	15	4.1	partir	12	12	7	1.9
prueba	23	19	8	1.5	probaron	12	12	6	1.3
<b>Adjectifs</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	probado	32	21	6	2.5
rico	49	22	3	5.7	estaba	29	18	4	5.8
bueno	201	66	2	5.9	utiliza	12	10	4	5.4
artesanal	4	4	2	4.0	prueban	13	10	4	2.6
dulce	26	12	2	6.3	dicen	39	20	3	4.8
<b>Verbes</b>	<b>Freq</b>	<b>CoFreq</b>	<b>Score</b>	<b>MDist</b>	tenía	24	14	3	3.0
hablar	46	46	28	2.3	pensó	9	7	3	3.4
elaborado	47	44	23	0.8	decían	12	8	2	2.8
utilizadas	31	31	19	4.4	han	36	16	2	3.8
definir	37	34	17	3.2	hacer	62	24	2	2.5
probó	38	33	15	2.0					

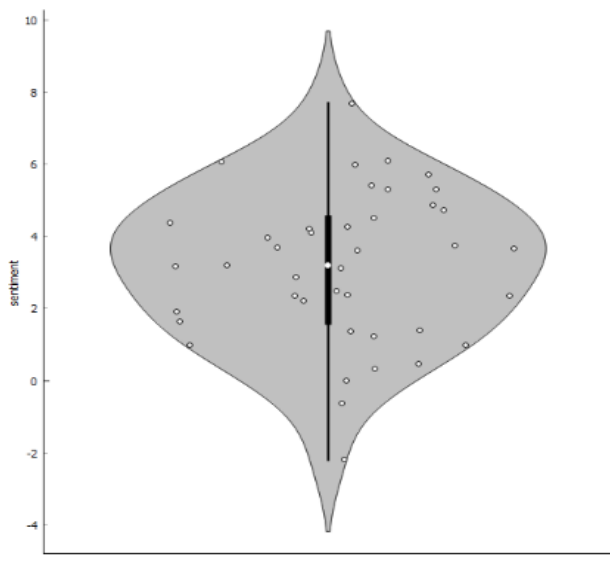
Tableau 21 : cooccurrence de chocolate dans le sous-corpus professionnel producteur de cacao (TXM)

Source : élaboré par nous



Cependant, les adjectifs cooccurrents *rico*, *bueno*, *artesenal* et *dulce* appartiennent au discours du producteur et représentent la grande majorité des termes qu'ils utilisent pour décrire les sensations que laisse le chocolat en bouche. Une étude plus approfondie du discours sensoriel des producteurs de cacao est présentée dans le Chapitre 4 : Empreinte du discours sensoriel dans les situations communicatives situées.

Selon l'analyse des sentiments (Sentiment Analysis) du programme Orange, les producteurs de cacao n'ont pas un discours aussi positif que les producteurs de chocolat. Avec une moyenne légèrement supérieure à 3, leur discours est à mi-chemin entre les résultats de l'analyse de sentiment du discours du corpus ambiant et ceux de l'analyse de sentiment du discours du corpus des producteurs de chocolat. Certains agriculteurs, comme Moisés Zambrano, soulignent la difficulté du travail au champ, mais la grande majorité est satisfaite de la rentabilité du produit vendu vantant les mérites de leur travail (*Figure 42*).



*Figure 42 : Représentation de l'analyse de sentiment des textes du sous-corpus professionnel producteur de cacao (Orange)*

*Source : élaboré par nous*

Même si la filière de production du chocolat adapte de nombreux formats issus de la filière vitivinicole, notamment dans ses stratégies de vente comme la roue des arômes

ou les notes de dégustation, le modèle de production du chocolat considéré comme produit haut de gamme n'est pas comparable au modèle de production vinicole tel qu'il est connu en France. En effet, il est difficilement concevable qu'un viticulteur français n'ait jamais goûté au vin produit à partir de son cépage. Un viticulteur dans son climat en Bourgogne ne laisserait pas le soin à un autre professionnel de produire et vendre le vin à partir du Pinot Noir qu'il cultive. En Équateur, le cacao *fino de aroma* est considéré comme un cacao noble, il est produit par des agriculteurs qui n'ont pas toujours eu l'occasion de goûter la tablette de chocolat plusieurs fois primée dans des concours internationaux qui est fabriquée à partir du fruit de leur travail. Cette réalité est le décor social de l'écosystème linguistique d'un pays avec de deux cents ans de tradition d'exportation de matière première débutant dans la transformation de sa *pepa de oro* pour exporter de l'orfèvrerie au lieu d'un simple minerai.

### 3.2.3. Moule discursif représentatif du discours professionnel du producteur de cacao en kichwa

Parmi les quarante et un agriculteurs qui ont accepté de participer aux entrevues en espagnol, trente d'entre eux n'ont pas souhaité réaliser l'entrevue en kichwa, onze ont accepté de répondre en kichwa auxquels il faut ajouter huit agriculteurs qui ont seulement accepté de répondre en kichwa. Le sous-corpus SCP cacao kichwa a été compilé à partir de ces dix-neuf entrevues (*Tableau 22*).

<b>sous-corpus SCP cacao kichwa</b>	<b>Token</b>	<b>Vocabulaire</b>	<b>Densité lexicale sur 1000 tokens aléatoires</b>
Nombre total	6416	1491	0.430

*Tableau 22 : Paramètres du sous-corpus professionnel producteur de cacao*  
*Source : élaboré par nous*

Les entrevues en kichwa ont été réalisées dans le même écosystème linguistique que les entrevues réalisées en espagnol. Les questions ont été posées par des natifs bilingues kichwa-espagnol ; et le discours en kichwa a été compilé à partir des transcriptions faites par ces mêmes natifs bilingues. De façon à pouvoir apprécier le contenu de ces entrevues, il a été possible de traduire huit d’entre elles en espagnol par un des locuteurs natifs bilingues qui a participé à la transcription. Considérant la taille réduite de ce corpus, il est reproduit en *Annexe 16* : Sous-corpus des producteurs de cacao en kichwa. De façon à comparer la production discursive des producteurs de cacao en espagnol et en kichwa, il est important de considérer le texte en espagnol qui correspond à la traduction du discours en kichwa et ainsi pouvoir analyser le volume discursif aussi bien en espagnol qu’en kichwa (*Tableau 23*).

	<b>Moyenne des termes</b>	<b>Moyenne des signes</b>
Discours produit en kichwa (8)	246,25	1873,63
Discours produit en kichwa traduit à l'espagnol (8)	342,50	1796,63
Discours produit en kichwa (11)	302,20	2068,80
Correspondance en espagnol (11)	420,32	1983,78
Discours produit en espagnol (11)	548,30	2736,60
Pourcentage de différence entre le discours produit en espagnol et en kichwa d’un même locuteur	30,45 %	37,95 %

*Tableau 23 : Comparaison de la production discursive en espagnol et en kichwa*  
*Source : élaboré par nous*

Pour cette analyse, il a été considéré la moyenne de termes des textes produits ainsi que la moyenne de signes de ces textes. Comme il est possible de le constater dans le *Tableau 23*, la moyenne du nombre de termes des huit textes produits en kichwa qui ont été traduits à l’espagnol est de 246,25 ; ce qui, selon la traduction, correspond à une moyenne de 342,50 termes en espagnol pour transmettre le même message, soit une augmentation de 40 %. Cependant, la moyenne du nombre de signes des huit textes produits en kichwa qui ont été traduits à l’espagnol est de 1873,63 ; ce qui, selon la

traduction, correspond à 1796,63 signes en moyenne en espagnol pour transmettre le même message, soit une diminution de 4%. Ce paradoxe s'explique par le fait que le kichwa soit une langue agglutinante qui présente aussi bien des flexions de cas nominales, adverbiales, que verbales. Pour illustrer cet exemple, la première question du guide d'entretien pour les entretiens en kichwa demande le nom de la personne : *ima shutita kanki* ? Ce qui littéralement serait en français : quel nom tu es ? (Quel est votre/ton nom ?) et en espagnol : *cuál nombre es/eres* ? (Cuál es su/tu nombre?) Dans le terme *shuti-ta*, il est possible d'observer le nom *shuti* et le suffixe *-ta* du contrastif (Itier, 2017 : 56), homonyme de l'accusatif dans la variété amazonienne du quichua, dont la fonction est de marquer et de clore le syntagme interrogatif *ima* (quoi). Avec cet exemple il est aisé de vérifier la différence de termes alors que le nombre de signes reste comparable. le mot *kanki* est le verbe 'être' (ka-) conjugué à la 2e p. (-nki). Le but de ce travail n'étant pas de proposer une recherche sur le kichwa de l'Amazonie équatorienne, pour de plus amples informations sur les caractéristiques de cette langue et ses variations, nous renvoyons aux ouvrages de Grzech (2016), de Mújica et Goldáraz (2010) et de Itier (2017). Les arguments présentés obligent le chercheur à considérer le nombre de signes et non le nombre de termes pour présenter la comparaison de l'étendue d'une production discursive entre espagnol et kichwa dans une situation de communication spécifique. A partir de cette réflexion, il est possible de conclure que le discours des locuteurs bilingues kichwa-espagnol qui ont accepté de réaliser l'entretien dans les deux langues, est de 38 % plus étendu en espagnol qu'en kichwa. Une des hypothèses pour expliquer cette différence est la supposition de Grzech sur le rapide déclin de la langue (2016 : 25). Mais il peut aussi être considéré que le kichwa s'utilise plus facilement dans un écosystème linguistique intime, alors que pour communiquer dans des écosystèmes linguistiques professionnels, les locuteurs kichwa utilisent plus naturellement l'espagnol.

Lors de l'analyse du corpus compilé en kichwa, la lemmatisation et l'étiquetage des parties du discours n'ont pu être réalisés par manque d'existence de fichiers de références pour le kichwa. En conséquence, les statistiques présentées se limitent aux

termes sans étiquetage spécifique. Le nuage de mot met en évidence le caractère agglutinant du kichwa. En effet, dans la (Figure 43), trois flexions de chocolat se retrouvent en premier plan : *chocolateta* (-ta = accusatif), *chocolate* (base substantive), *chocolatemanta* (-manta = ablatif) ; ainsi que cinq flexions de cacao : *cacao* (base substantive), *cacaota* (-ta = accusatif), *cacaowa* (-wa = commitatif « avec »), *cacaokunata* (-kuna = pluriel, -ta = accusatif), *cacaomanta* (-manta = ablatif) (Itier, 2017 : 50). Comme précédemment observée lors de l’analyse des sous-corpus professionnels en espagnol, la nature même de l’entrevue positionne les termes utilisés pour poser les questions dans la liste des termes les plus fréquents, par exemple le pronom interrogatif *ima* (que/quoi) et ses flexions : *imasna* (combien), *imasna pachata* (quand), *imarashpata* (pourquoi).



Figure 43 : Nuage des cent premiers termes du sous-corpus professionnel producteur de cacao en kichwa (Orange)  
Source : élaboré par nous

L'extraction des termes les plus utilisés retourne les mots des questions avec les indices les plus élevés (*Tableau 24*). Cependant, une fois retiré les noms propres ; après une analyse plus précise, il est possible d'observer la présence de déictiques comme *kay* et *chay* (adjectifs démonstratifs) dont l'usage est en accord avec le fait que les entrevues soient réalisées dans les plantations de cacao. Les termes *ñukanchi* (nous) et *ayllu* (famille) sont en accord avec l'organisation familiale de l'exploitation. Dans le discours en kichwa, les références au système d'agriculture diversifiée sont aussi présentes : *palanta* (banane plantain), *tarpusha* (semence), *llaktamanta* (de la terre, du peuple), *lumu* (manioc), *tarpu* (la semence), *muyukunata* (des graines, accusatif). Bienqu'il existe un terme kichua pour le cacao sylvestre : *kampik*, celui-ci n'est pas utilisé, car le cacao cultivé est une espèce modifiée mise à disposition par des entreprises soutenues par le ministère de l'agriculture. Le terme *cacao* est ainsi inclus et décliné dans le discours kichwa tout comme *chocolate* qui est le produit fabriqué à partir de ce cacao, mais qui ne fait pas partie de leur réalité. En effet, pour répondre aux questions sur le goût du chocolat fabriqué à partir du cacao produit, les participants aux entrevues font référence à la boisson chocolatée (*upina*) préparée artisanalement à partir du cacao séché, toasté et moulu comme l'illustre cet exemple dont la traduction réalisée par l'intervieweur se trouve à la suite de la contribution entre parenthèses :

*C-014 : entrevistador: rimanata ushankichu imasami chocolateta llukshin kampak cacao rurakpi? (¿puede definir las características del chocolate elaborado con su cacao?<sup>52</sup>)*

*C-015 : entrevistado: kay cacaomanta chocolate manta llukshinmi gustu asnak y mishki chocolateta mashtinkawa cacaota pitinchi chiwasha apamusha chakichinchi intiy cuatro días chakichishka washa chakishkay chiwasha tustanchi tustasha kutasha kutashkawasha churanchi ansa azúcar vainilla masanchi chipi chiwasha yanunchi y gustu upina tukun (de es este cacao lo elaboramos chocolates de*

---

<sup>52</sup> Pouvez-vous définir les caractéristiques du chocolat élaboré à partir de votre cacao ?

*calidad y para la elaboración del chocolate lo cortamos el cacao después lo secamos en el sol por cuatro días una vez secado lo tostamos luego lo molemos y se agraga un poco de azúcar y vainilla así es como lo preparamos para tomar<sup>53</sup>)*

*C-016 : ima yachikta charin kay chocolateta chasnallata imata yuyachin? (¿qué sabor tiene este chocolate y a qué le hace pensar?<sup>54</sup>)*

*C-017 : entrevistado: chocolateka gustumi wawakunata yanusha upichinkawa vainillata churakpi gustu mishki upina tukun shinallata churanchi chocolatey agengibrewa mezclanchi shinallata gustu llukshin chitami wawakunata tutamantawna yanusha upichinchi (entrevistado: el chocolate es bueno para dar de tomar a los niños a lo que agregamos vainilla tiene un aroma delicioso así mismo lo agregamos ajengibre y lo mezclamos<sup>55</sup>)*

Dans le discours produit en kichwa, il est facile d'identifier certains emprunts faits à l'espagnol. Étant donné que le kichwa ne possède que trois voyelles ; i, u et a ; les termes utilisant le e et le o sont importés de l'espagnol comme : *chocolateta*, *cacaomanta*, *agengibrewa*, *mezclanchi*. D'autres emprunts sont plus difficiles à identifier à cause des modifications faites pour les intégrer dans le discours. Le terme *gusto* est utilisé et décliné en kichwa sous la forme de *gustu*. Cet emprunt peut s'expliquer par le fait que *gustar* se traduit en kichwa par *allikachina*, ce terme est composé de *alli* qui traduit bon, suivie de deux suffixes de dérivation verbale et de l'infinitif, *allikachina* signifie 'trouver bon'. Il est aussi possible d'observer un autre ensemble de termes associés au discours sensoriel : *alli* (bon, bien), *asnak* (odorant), *gustu* (goût), *kamana* (goûter), *mishki* (doux, miel, savoureux, délicieux, exquis), *sumak* (beau, belle), *yachik* (saveur).

---

<sup>53</sup> *de ce cacao, nous élaborons des chocolats de qualité et pour l'élaboration du chocolat nous coupons le cacao après nous le séchons au soleil pendant quatre jours une fois séché nous le toastons après nous le moulons et on ajoute un peu de sucre et de vanille c'est comme ça que nous le préparons pour le boire*

<sup>54</sup> *Quelle est la saveur de ce chocolat et à quoi vous fait-il penser ?*

<sup>55</sup> *le chocolat est bon pour donner à boire aux enfants quand nous y ajoutons la vanille il a un arôme délicieux de la même manière nous ajoutons du gingembre et nous le mélangeons*

Ces exemples mettent en évidence que, comme dans le discours en espagnol, les termes utilisés pour décrire les sensations que laisse la dégustation du chocolat restent des termes qui appartiennent au langage commun sans qu'ils puissent être qualifiés de descriptifs sensoriels spécifiques du cacao ou du chocolat.

Freq	Termes	Trad	Freq	Termes	Trad
44	<i>alli</i>	bon, bien	22	<i>llaktamanta</i>	de la terre, du peuple
19	<i>allimi</i>		24	<i>llankaypi</i>	travail
4	<i>allita</i>		43	<i>llankayta</i>	
28	<i>ansa</i>	peu	33	<i>llukshin</i>	sortir
13	<i>aparkanki</i>	Produire	15	<i>lumu</i>	manioc
16	<i>apawki</i>		53	<i>mana</i>	négation
5	<i>apawnki</i>		42	<i>maykan</i>	quel, quelle
5	<i>asha</i>	beaucoup	19	<i>mikuna</i>	manger
5	<i>ashakata</i>		38	<i>mishki</i>	doux, savoureux
20	<i>ashka</i>		15	<i>muyukunata</i>	des fèves (accusatif)
21	<i>ashkata</i>		45	<i>nisha</i>	dit
13	<i>asnak</i>	odorant	91	<i>ñuka</i>	je, moi
28	<i>ayllu</i>	famille	127	<i>ñukanchi</i>	nous, nôtre
122	<i>cacao</i>	cacao	19	<i>paktaykunata</i>	responsabilités (accusatif)
10	<i>cacaokunata</i>		21	<i>palanta</i>	banane plantain
9	<i>cacaomanta</i>		19	<i>rimanata</i>	parler
30	<i>cacaota</i>		27	<i>rimanawn</i>	
17	<i>cacaowa</i>		23	<i>rimarin</i>	
26	<i>charin</i>	avoir	30	<i>runakuna</i>	personnes
10	<i>charinchi</i>		21	<i>rurakpi</i>	faire
18	<i>charini</i>		19	<i>ruranki</i>	
20	<i>charinki</i>		19	<i>rurayta</i>	
16	<i>chasnallata</i>	ainsi	12	<i>sabor</i>	saveur (mot espagnol)
11	<i>chay</i>	adjectif démonstratif éloigné	18	<i>shamunki</i>	venir
63	<i>chocolate</i>	chocolat	20	<i>shimikunata</i>	langages (accusatif)
30	<i>chocolatemanta</i>		13	<i>shina</i>	comme
105	<i>chocolateta</i>		36	<i>shinallata</i>	également
10	<i>chunka</i>	dix	12	<i>shuti</i>	nom
67	<i>gustu</i>	goût	16	<i>shutita</i>	



98	<i>ima</i>	Pronoms interrogatifs	12	<i>shutikunamanta</i>	personnel
15	<i>imarashpata</i>		37	<i>sumak</i>	beau, belle
18	<i>imasami</i>		12	<i>tarpani</i>	semer
32	<i>Imasna (pachata)</i>		13	<i>tarpasha</i>	
16	<i>imasnata</i>		18	<i>tarpusha</i>	
44	<i>imata</i>		18	<i>tukuy</i>	tout, entier
30	<i>kamakpi</i>	goûter	11	<i>upina</i>	boire, boisson
14	<i>kamashpa</i>		18	<i>ushankichu</i>	pouvoir
42	<i>kampak</i>	ton	10	<i>vainilla</i>	vanille
58	<i>kan</i>	il y a	13	<i>wata</i>	année
10	<i>kani</i>	je suis	20	<i>yachikta</i>	saveur (accusatif)
19	<i>kanki</i>	tu es	34	<i>yachin</i>	savoir, sensation
255	<i>kay</i>	adjectif démonstratif proche	19	<i>yuyachin</i>	penser, réfléchir
13	<i>kaypi</i>				

Tableau 24 : Mots les plus utilisés du sous-corpus professionnel producteur de cacao en kichwa (Orange)

Source : élaboré par nous

Le regroupement (Figure 45) a été élaboré à partir des termes du Tableau 24. Le premier regroupement ne laisse que peu de place à l'interprétation : *ñuka* (je), *lumu* (manioc), *apawna* (produire) confirmant que d'autres produits sont cultivés avec le cacao. Les termes *ansa* (peu), *gustu* (goût) et *mana* (négation) partagent le deuxième regroupement confirmant que peu de participants ont goûté au chocolat fabriqué à partir du cacao qu'ils produisent, .

Dans le regroupement C3, les termes associent la qualité, la saveur, à l'expérience familiale qui prépare un chocolat, la boisson, en suivant leur propre processus de fabrication en relation avec le regroupement C4 qui associe ce processus à la vanille. Il est important de savoir que la vanille est un produit abondant à l'état naturel en Amazonie, tout comme la cannelle. Ces deux ingrédients se retrouvent dans la grande majorité des cas présents dans la boisson chocolatée offerte aussi bien aux touristes qu'aux résidents. Les regroupements C5, C6 et C7 sont en étroite relation avec la question de savoir pourquoi ils cultivent du cacao ; question à laquelle ils répondent qu'il sème du cacao pour pouvoir s'en sortir. La raison est directement liée au rendement

économique de la vente du cacao. Le huitième regroupement sont les différentes flexions de *bon / bien* qui associés au neuvième regroupement répond à la question sur les responsabilités qu'ils ont dans leur travail : ils sont responsables des fèves de cacao qui viennent de la terre pour faire un bon chocolat. Le dernier regroupement confirme que le cacao « odorant » (asnak) qui sert à fabriquer la boisson chocolatée est produit au milieu d'autres produits comme la banane plantain.

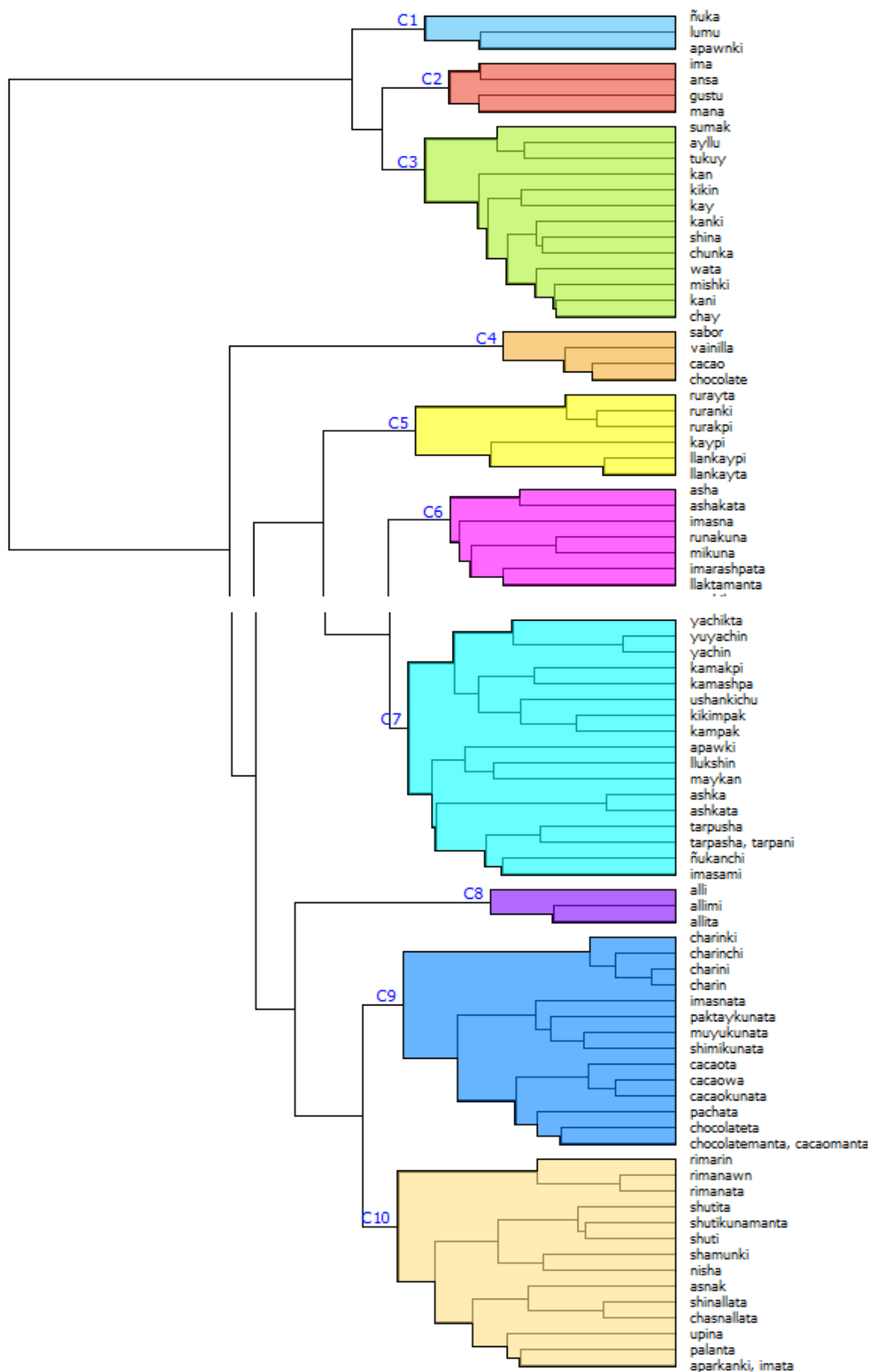


Figure 44 : Regroupement hiérarchique des mots fréquents du sous-corpus professionnel producteur de cacao (Orange)

*Source : élaboré par nous*

À la différence du discours produit en espagnol, les cultivateurs de cacao ne s'emploient pas qu'à fournir une fève répondant aux standards de qualité du producteur de chocolat, ils travaillent aussi pour la qualité de la boisson chocolatée qu'ils élaborent eux-mêmes à partir de la fève de cacao et d'ingrédients ajoutés comme la vanille. Cette boisson est souvent proposée avec des *muchines* préparés à base de manioc et de fromage (*Figure 45*).



*Figure 45 : Photo d'une boisson chocolatée avec son muchin prise à la communauté « Porvenir »*

*Source : élaboré par nous*

La présence des différentes flexions augmente la densité des liens de la projection bidimensionnelle (*Figure 46*). Cependant, il est possible d'observer que les regroupements C1, C5, C6 et C7 sont les plus souvent en relation.

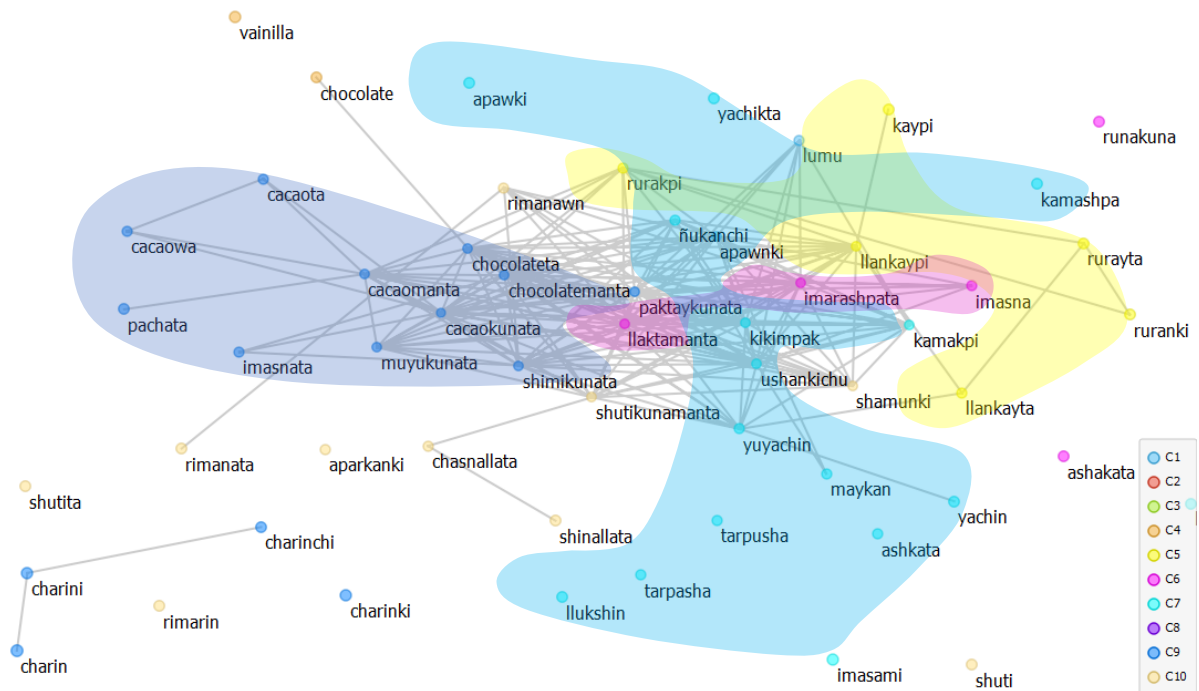


Figure 46 : Agrandissement de la projection bidimensionnelle (MDS) des mots les plus utilisés du sous-corpus producteur de cacao (Orange)

Source : élaboré par nous

Ces liens confirment que les volontaires qui ont répondu au questionnaire ont présenté plus de facilités pour s'exprimer sur leurs responsabilités dans leur travail, sur leurs activités au champ, que sur les impressions que peut laisser la dégustation faite d'un chocolat fabriqué à partir de leur fève de cacao. Leur discours ainsi compilé et analysé démontre que la tablette de chocolat ne fait pas partie de leur écosystème linguistique, pas plus que la roue des arômes, instaurant que la boisson chocolatée est la seule forme de consommation de la fève de cacao que les communautés pratiquent.

La représentation graphique du réseau sémantique des termes dont la fréquence est supérieure à 25, dans une fenêtre de deux mots à gauche et à droite, des cooccurrences dont la fréquence est supérieure à 10 (Figure 47) met en évidence les termes utilisés pour les questions, surlignés en jaune, et deux termes utilisés pour décrire le goût du chocolat : *mishki* (doux, savoureux), *sumak* (bon, bien)

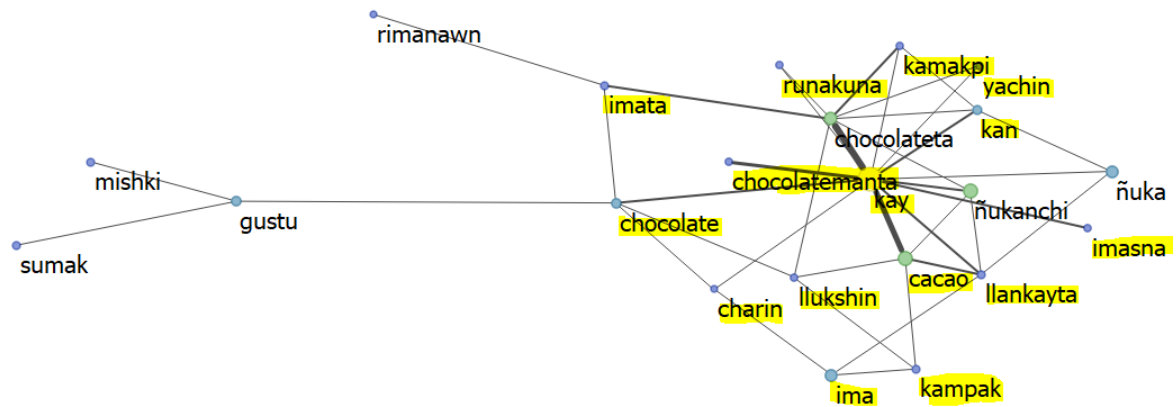


Figure 47 : Réseau sémantique des principales cooccurrences du sous-corpus professionnel producteur de cacao (Orange)

Source : élaboré par nous

Nombreux sont les paramètres étudiés qui mettent en évidence l'absence de la tablette de chocolat dans l'écosystème linguistique des producteurs de cacao aussi bien quand ils s'expriment en espagnol comme en kichwa. Dans leurs représentations sociales et leurs traditions, le chocolat est une boisson douce vanillée préparée artisanalement.

Ce troisième chapitre nous aura permis de continuer dans l'analyse approfondie des caractéristiques discursives des sous-corpus précédemment décrits de façon à pouvoir lire et interpréter les empreintes, ou l'absence d'empreintes, du discours sensoriel dans le sous-corpus ambiant et le sous-corpus des professionnels. La description de l'application d'outils textométriques pour analyser les sous-corpus compilés, y compris des caractéristiques telles que les occurrences de mots, les concordances et l'analyse des cooccurrences qui informent sur la fréquence, la distribution et l'utilisation sémantique d'un mot ou d'une expression fournissant ainsi des éléments quantitatifs pour interpréter les modèles sémantiques et syntaxiques. Quant à la méthode de classification par regroupement hiérarchique, elle mesure la dissemblance entre les individus d'un groupe et s'utilise pour comprendre l'organisation des données.

L'objectif de ce chapitre est d'utiliser une méthodologie d'exploration analytique des données sur un corpus représentatif afin de décrire les modèles discursifs de chaque

sous-corpus et d'analyser les stratégies utilisées par les professionnels et les *aficionados* pour exprimer leurs émotions et leurs passions liées au cacao ou au chocolat.

L'analyse du sous-corpus ambiant relatif au cacao et au chocolat identifie deux groupes principaux de termes : l'un lié à la qualité et à la saveur du cacao et du chocolat, et l'autre lié au processus de production et d'exportation. Les mots-clés peuvent être regroupés en différents thèmes, tels que la production, le patrimoine et les références spatio-temporelles. Il semble important pour le professionnel de la filière de promouvoir les aspects techniques de la production de cacao et de la présenter comme faisant partie du patrimoine équatorien sur les marchés internationaux. Cependant, l'analyse met en évidence l'absence d'un discours sensoriel : les cooccurrences des termes "cacao" et "chocolat" présentent des associations descriptives plutôt que sensorielles ou même émotionnelles.

Cette absence de discours sensoriel se retrouve dans l'analyse du sous-corpus relatif à la production et à l'exportation de chocolat qui met en évidence dans le discours du producteur la présence du processus de production, des stratégies de commercialisation, et de la qualité du cacao équatorien. Cependant, bien que les producteurs de chocolat utilisent une terminologie standardisée relative à la forme physique et au descriptif sensoriel du chocolat, ils ne produisent pas pour autant un discours spécifique pour exprimer l'émotion créée par le chocolat. Dans le discours des producteurs de cacao apparaissent des adjectifs communs pour exprimer les sensations que le chocolat laisse en bouche même si leur priorité discursive reste centrée sur la rentabilité du produit et les mérites de leur travail, le soutien économique étant la principale motivation des agriculteurs. Quand les producteurs de cacao acceptent de s'exprimer en kichwa, leur langue natale, ils confirment leur statut d'agriculteurs qui consomment une boisson chocolatée à la vanille, la plaquette de chocolat ne faisant pas partie de leur écosystème sociolinguistique. Dans leurs représentations sociales et leurs traditions, le chocolat est une boisson vanillée sucrée faite à la main.

## Chapitre 4 : Empreinte du discours sensoriel dans les situations communicatives situées de l'*aficionado*

---

### 4.1. Stratégies discursives de l'*aficionado*

Le processus de mise en évidence des stratégies discursives de l'*aficionado* repose sur la compilation du sous-corpus questionnaire (*Annexe 17* : sous-corpus questionnaire sans stimulus et *Annexe 18* : sous-corpus questionnaire avec stimulus) dont les caractéristiques sont présentées dans le *Tableau 2*. Il convient de préciser que pour l'analyse, le questionnaire est divisé en deux parties : les questions du recto de la feuille, réalisées sans stimulus, et les questions du verso de la feuille, posées après avoir proposé la dégustation d'un morceau de chocolat dont la teneur en cacao était soit de 70 % soit de 100 %. La première partie « permet d'accéder aux représentations mentales de[s] répondants sur ce produit » (Mancebot, 2019 : 72). Le verso, la seconde partie du questionnaire, permet d'accéder à un discours dont « la mise en mots est couplée et déclenchée par le sensoriel » (Mancebot, 2019 : 72). Après l'analyse discursive des deux parties, la comparaison des résultats met en évidence les similitudes et différences entre un discours déclenché par un processus cognitif et un discours déclenché par un processus sensitif.

Afin de caractériser les *aficionados* qui ont volontairement répondu au questionnaire et déterminer le public représenté, les réponses aux questions 1, 2, 3, 4, 5 et 9 ont été compilées dans le *tableau 25*

Le profil représentatif est indépendant du genre. D'un niveau d'étude supérieur, entre 30 et 50 ans, l'*aficionado* consomme régulièrement du chocolat (au moins une fois par semaine pour 60% des participants) ; et il présente une nette préférence pour le chocolat dont la concentration de cacao est de 70%.



<b>Origine</b>	<i>Equatorien</i>	<b>Niveau d'instruction</b>	<i>Secondaire : 8%</i>
<b>Genre</b>	<i>Femme :52%</i> <i>Homme :48%</i>		<i>Premier cycle universitaire : 69%</i> <i>Deuxième cycle universitaire : 23%</i>
<b>âge</b>	<i>18/22 : 2.3%</i>	<b>Consommation de chocolat</b>	<i>1 fois à l'année : 0.8%</i>
	<i>23/30 :18.5%</i>		<i>De 2 à 10 fois par an : 10.8%</i>
	<i>31/36 :28.5%</i>		<i>De 1 à 3 fois par mois : 27.7%</i>
	<i>37/41 :11.5%</i>		<i>1 fois à la semaine : 33.1%</i>
	<i>42/50 :21.5%</i> <i>+50 :17.7%</i>		<i>+ d'une fois à la semaine : 27.7%</i>
<b>Evaluation échantillon à 70 %:</b> 8,24/10		<b>Evaluation échantillon à 100 %:</b> 6,55/10	

Tableau 25 : Profil de l'aficionado

Source : élaboré par nous

## 4.2. Réponses au questionnaire sans stimulus

La nature – questionnaire ouvert – des textes constitutifs du sous-corpus représentatif du discours de l'aficionado limite le nombre de *token* aux réponses des participants, il est cependant remarquable que la densité lexicale soit comparable à celle du discours professionnel du producteur de chocolat (Tableau 10 et Tableau 26) .

sous-corpus SCQ sans stimulus	<i>Token</i>	Vocabulaire	Densité lexicale sur 1000 tokens aléatoires
Nombre total	2870	723	0.352

Tableau 26 : Paramètres du sous-corpus de l'Aficionado sans stimulus

Source : élaboré par nous

Cette similitude semble être une similitude de grand intérêt étant donné que l'intention du professionnel est de convaincre, à travers son discours, pour vendre son produit ; alors que celle de l'aficionado est de rechercher l'émotion qui lui fera acheter tel ou tel chocolat. Dans la filière du chocolat, ces deux locuteurs sont donc amenés à

communiquer régulièrement, il est alors légitime de se demander si leurs discours sur le chocolat sont congruents et en harmonie tout au moins en ce qui concerne le discours sensoriel produit.

Comme le producteur de chocolat, quand il est amené à parler du produit, l'*aficionado* prononce plus fréquemment le terme chocolat que cacao. Cependant, dans un discours où les questions sont directement orientées sur le chocolat, le concept de cacao reste fortement présent et semble indissociable du produit fini chez l'*aficionado* équatorien (Tableau 27).

<b>[espos="NC"]:</b> eslemma / F							
chocolate	70	pureza	5	familia	3	trabajo	2
sabor	43	cultivo	4	fruta	3	época	2
aroma	37	leche	4	nivel	3	esencia	2
cacao	33	materia	4	placer	3	fiesta	2
producto	25	proceso	4	tierra	3	gusto	2
calidad	23	sensación	4	vez	3	hijo	2
mundo	23	textura	4	abuelita	2	industrialización	2
dulce	15	variedad	4	alimento	2	infancia	2
golosina	12	amor	3	boca	2	pastel	2
país	8	azúcar	3	bombón	2	persona	2
barra	6	característica	3	casa	2	postre	2
delicia	6	contenido	3	comparación	2	producción	2
color	5	cremoso	3	componente	2	resto	2
exportación	5	energía	3	costa	2	satisfacción	2
olor	5	enfermedad	3	taza	2	tarde	2
paladar	5	esposo	3	tiempo	2		
<b>[espos="ADJ"]:</b> eslemma / F							
bueno	65	agradable	7	exquisito	3	especial	2
dulce	29	sabroso	7	intenso	3	excelente	2
rico	23	caliente	6	nutricional	3	fuerte	2

ecuatoriano	23	orgánico	6	propio	3	gran	2
delicioso	22	puro	5	sano	3	inigualable	2
único	15	fino	4	suave	3	oscuro	2
amargo	12	económico	3	aromático	2	original	2
nuestro	10	exótico	3	energético	2	particular	2
<b>[espos="V.*"] : eslemma / F</b>							
ser	85	probar	7	poder	4	faltar	2
tener	16	gustar	6	dar	3	llevar	2
hacer	14	consumir	5	decir	3	obtener	2
pensar	12	deber	5	producir	3	permitir	2
caracterizar	9	haber	5	realizar	3	procesar	2
elaborar	8	sentir	5	compartir	2	servir	2
<b>[espos="ADV"] : eslemma / F</b>							
muy	28	bien	4	poco	3	ya	3

Tableau 27 : Lemmes de fréquence supérieure à 1 du sous-corpus *Aficionado sans stimulus (TXM)*

Source : élaboré par nous

Il serait intéressant de réaliser ce questionnaire dans un pays non producteur de cacao, qui présente donc un écosystème linguistique différent, pour vérifier si l'*aficionado* conserve cette relation conceptuelle entre chocolat et cacao. Contrairement au producteur qui focalise son discours sur une thématique de processus de production et de commercialisation, l'*aficionado* centre son discours sur le produit : *producto*<sup>56</sup>, ces caractéristiques : *sabor, aroma, calidad, barra, color, olor, pureza, materia, textura, variedad, característica, cremoso, alimento, esencia, dureza, viscosidad*, tout en répondant aux questions posées qui cherchent à savoir ce que le terme chocolat évoque chez l'*aficionado* : *dulce (caramel), golosina, energía, esposo, familia, abuelita, bombón, casa, época, fiesta, hijo, infancia, pastel, postre, tarde, taza, amigo, cultura, descanso, manjar, niñez, privilegio*. L'*aficionado* n'oublie pas dans son discours de faire

<sup>56</sup> Les concordances des termes cités sont consultables en *annexe 17*; *Error! No se encuentra el origen de la referencia.*

référence aux sensations qu'évoque le simple terme chocolat : *delicia, paladar, sensación, amor, placer, boca, gusto, satisfacción, afecto, alegría, calor, enamoramiento, exquisitez, felicidad, gana, relajamiento, romanticismo, ternura, tranquilidad, vida*. Les autres noms communs sont relatifs à l'origine : *país, costa*, à des expressions superlatives : *mejor del mundo, explosión de sabores, la delicia de los sentidos*, un chocolat digne d'être exporté (*exportación*), associé aux ingrédients suivants : *leche, azúcar, fruta, almendras* ; et pour terminer, comme son interlocuteur le professionnel, il fait aussi référence au processus sans lui accorder autant d'importance : *cultivo, industrialización, producción, trabajo, actividad, agricultor, comercio, cosecha*.

Sans être absolument spécifique d'un discours du sensoriel sur le chocolat, les adjectifs utilisés sont plus liés aux sens que ceux des producteurs de chocolat : *bueno, dulce, rico, delicioso, amargo, agradable, sabroso, exótico, exquisito, intenso, suave, aromático, excelente, fuerte*. Grâce à ces caractéristiques sensorielles, l'*aficionado* qualifie le chocolat équatorien comme : *único, puro, fino, sano, energético, especial, gran, inigualable, original, particular*. Il n'oublie pas de rappeler que ce produit d'intérêt économique est de grande qualité, car il est : *ecuatoriano, nuestro, propio, orgánico, sano*. Contrairement au discours du producteur, les verbes ne présentent pas de références spécifiques au ressenti, il semble que la stratégie discursive de l'*aficionado* soit plus centrée sur la qualification que sur l'action. *Ser, tener* et *caracterizar* sont des verbes qui décrivent les caractéristiques du produit ; *hacer* et *dar* évoquent un résultat : *hace pensar, hace sentir, da energía* ; *pensar* est une reprise du verbe principal de la question posée. Le verbe *elaborar* permet de faire référence au produit de base : le cacao. *Probar* associé à *deber* exprime une injonction pour inviter les gens à goûter le produit ou encore à exprimer une nécessité : *que nunca debe faltar*. *Gustar* et *sentir* sont directement associés au discours sensoriel, sensation qui n'est ici qu'évoquée : *me gusta deleitar chocolate, me gusta el chocolate ecuatoriano, hace sentir satisfecho y animado, hace sentir ganas*. Les adverbes ne présentent pas de caractéristiques remarquables.



La mise en évidence des mots-clés selon la méthodologie YAKE ! (Campos et coll., 2020) appliqué à l'espagnol<sup>57</sup>, permet de souligner les termes dont la fréquence est particulièrement remarquable dans cet écosystème linguistique (*tableau 28*).

41 Keywords Q sans					
Word	YAKE!	Word	YAKE!	Word	YAKE!
leche	0,04	caliente	0,03	chocolate	0,01
fino	0,04	aroma	0,02	golosina	0,01
relajante	0,04	sabor	0,02	económico	0,01
felicidad	0,04	amargo	0,02	orgánico	0,01
país	0,04	ecuador	0,02	calidad	0,01
amor	0,03	barra	0,02	excelente	0,01
olor	0,03	único	0,02	saludable	0,01
chocolates	0,03	alegría	0,02	mundo	0,01
delicioso	0,03	placer	0,02	aromático	0,01
delicia	0,03	dulce	0,02	exquisito	0,01
producto	0,03	caracteriza	0,02	puro	0,01
pienso	0,03	ecuatoriano	0,01	agradable	0,01
rico	0,03	cacao	0,01	sabroso	0,01
pureza	0,03	pensar	0,01		

Tableau 28 : 41 mots-clés du sous-corpus Aficionado sans stimulus (Orange)

Source : élaboré par nous

Le premier terme *leche* fait référence à la boisson : *acompañado de un vaso de leche, chocolate en leche, chocolate caliente con leche, en la tarde con leche* ; illustrant une forme traditionnelle de consommer le chocolat en Équateur. Le terme *fino* est aussi bien utilisé pour décrire la délicatesse du chocolat : *es muy fino, auténtico fino dulce, fino único* ; que pour faire référence au type de cacao : *original chocolate fino de aroma, como cacao fino de aroma, el mejor fino de aroma*. Il est ensuite remarquable de voir que les évocations émotives ont une bonne représentation dans les mots-clés : *relajante, felicidad, amor, delicia, pureza, alegría, placer*.

<sup>57</sup> Algorithme intégré dans le programme Orange

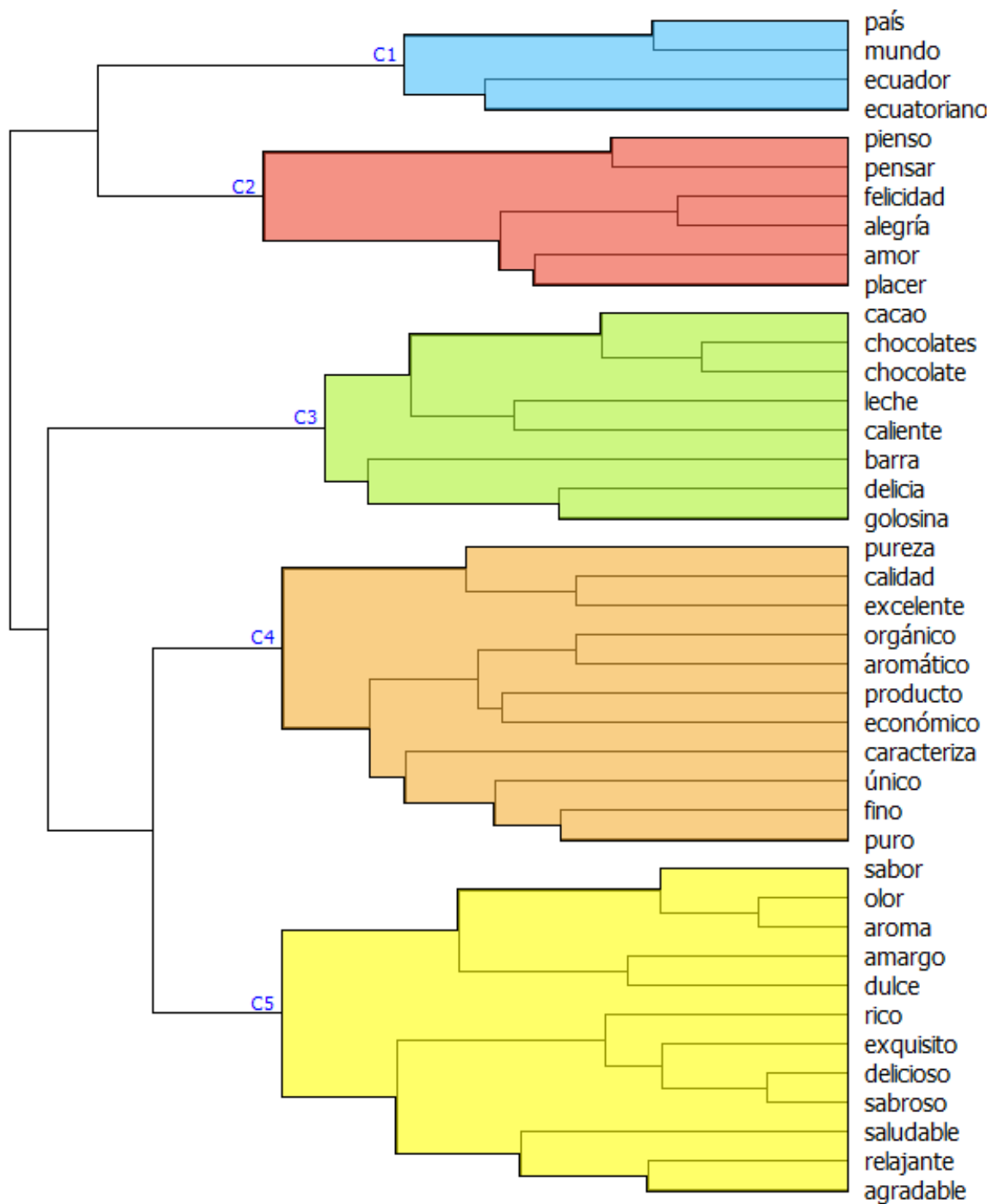
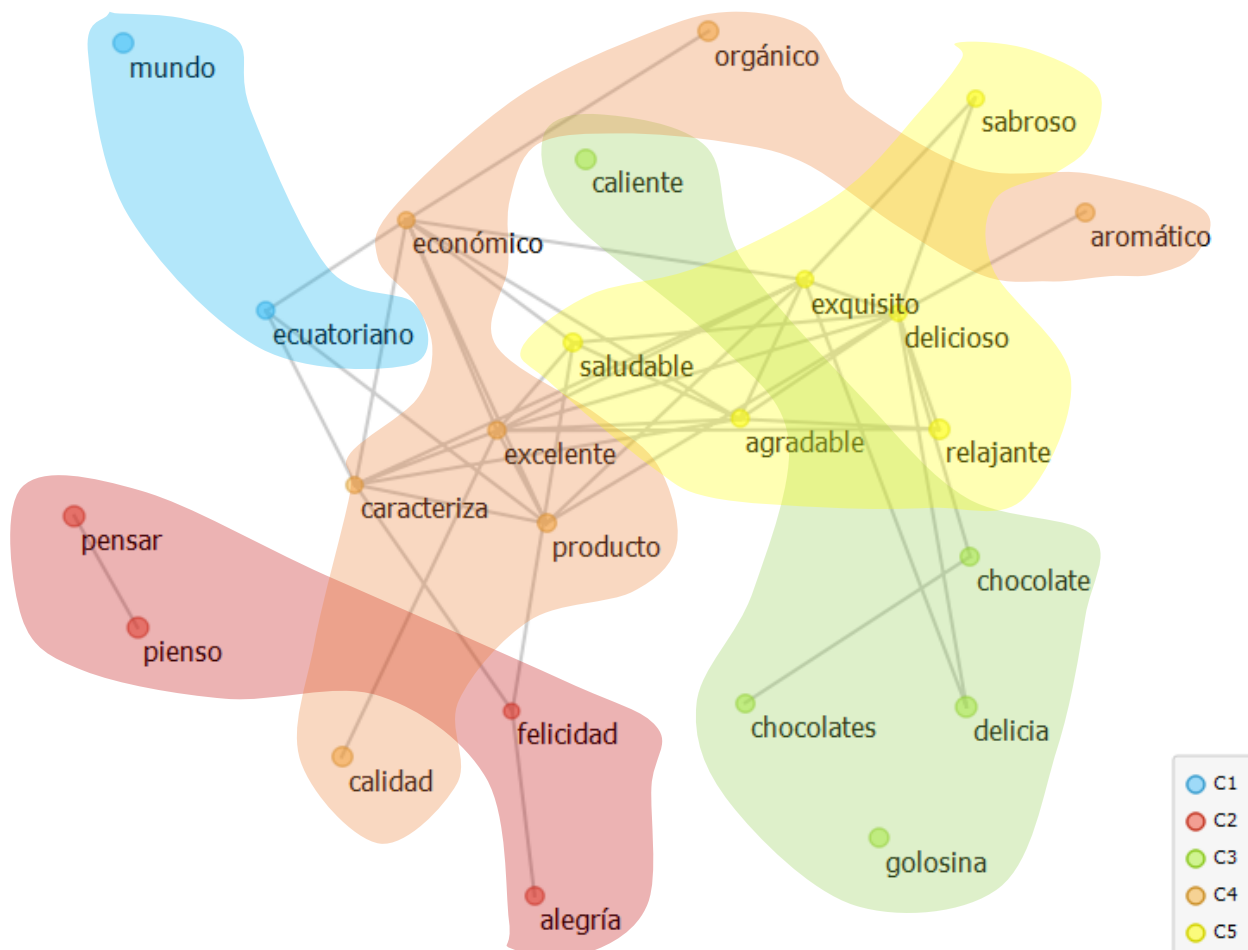


Figure 49 : Regroupement hiérarchique des mots-clés du sous-corpus *Aficionado sin stimulus* (Orange)

Source : élaboré par nous

Le regroupement hiérarchique (*Figure 49*) de ces mots-clés confirme que dans les représentations sociales équatoriennes, le concept de chocolat est considéré comme une source de joie et de plaisir (C2) décrite à partir d'expressions hédoniques (C5), d'orgueil national (C1), et de caractéristiques exceptionnelles (C4). Le regroupement C3 faisant référence aux formes de consommation du chocolat, en tablette ou en boisson à base de lait. Il est ici important de se rappeler que ce discours est produit à partir des représentations mentales évoquées par les questions du questionnaire sur le chocolat. Cette mise en mot est le résultat d'un processus cognitif qui se nourrit aussi bien de l'écosystème linguistique du moment – le salon du chocolat à Quito – que des souvenirs évoqués et construits par les expériences personnelles que chaque participant peut associer au concept de chocolat.



*Figure 50 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus Aficionado sin stimulus (Orange)*



Source : élaboré par nous

La fonction d'échelle multidimensionnelle (Borg et coll., 2018), *multidimensional scaling (MDS)* met en évidence le lien central entre C4 et C5 (*Figure 50*) où se mêlent les sensations comme : *relajante*, les caractéristiques comme : *saludable* et les expressions hédoniques comme *exquisito*. Non seulement le discours sensoriel de l'Aficionado exploite plusieurs stratégies, mais il occupe aussi une place centrale dans son discours sur le produit, ce qui contraste avec tous les corpus précédemment analysés. En périphérie de ces deux regroupements s'observent les notions de joie, d'origine et d'un chocolat considéré comme une gourmandise.

La taille du corpus étudié limite l'utilisation des cooccurrences pour définir le réseau sémantique d'un terme. Cependant, l'outil *Network explorer* du programme Orange offre une représentation graphique des termes dont la fréquence de cooccurrence est supérieure à 5 dans une fenêtre de quatre mots à gauche et à droite, des termes dont la fréquence est supérieure à 10 (*Figure 51*).

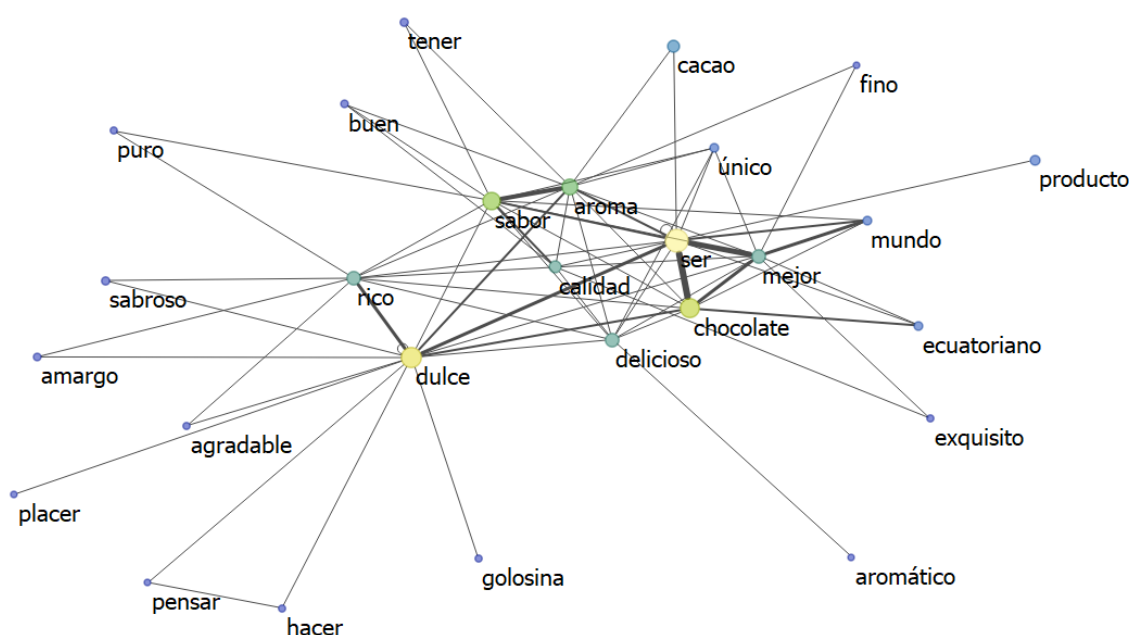


Figure 51 : Réseau sémantique des principales cooccurrences du sous-corpus Aficionado sans stimulus (Orange)

Source : élaboré par nous

Selon ce réseau sémantique, pour les *aficionados* participant au questionnaire, le chocolat équatorien est le meilleur du monde. Il est exquis, délicieux, unique, pur. Il fait penser à une agréable gourmandise riche et savoureuse. Les expressions particulières à chaque locuteur en réponse à la question six (*Si digo “chocolate”, ¿A qué le hace pensar?*) illustrent ce sentiment général : *es la delicia de los sentidos, pienso en una taza de chocolate caliente en la tarde, lleno de ternura que hace sentir satisfecho y animado, sabor delicioso que se derrite en la boca provocando una sensación satisfactoria, hace referencia a una persona amable detallista y con buenos sentimientos, recuerdos positivos, algo dulce que se deshace en nuestro paladar, amistad familia tranquilidad paz, recordar el chocolate de la abuelita, relajamiento dulce, saborear la vida, a sentirse bien consigo mismo, pensar en chocolate es sentirse bien en la infancia, momentos vividos de la niñez, época de san valentin, navidad, cumpleaños, amor hacia mi esposa e hija, al escuchar chocolate me hace pensar en romanticismo algo delicioso, se hace agua la boca, se consume para alegrar el día, amor me recuerda a mi hijo a mi infancia*. Cependant, ces exemples mettent en évidence que plus qu’une stratégie terminologique qui en un mot exprimerait une sensation, l’*aficionado* met en œuvre une stratégie discursive autour d’expressions évocatrices d’actuations sociales riches en émotions qui font appel à la propre expérience de celui qui pose la question. Il semble que le participant cherche une résonance chez celui qui pose la question pour partager ses émotions plus que pour les décrire.

Les réponses à la septième question (*Si debe explicar a un amigo que caracteriza el chocolate ecuatoriano, ¿qué diría?*) sont directement liées à l’orgueil national de disposer d’un tel produit : *es único y mejor del mundo por su calidad y pureza, el mejor chocolate del mundo, el cacao que se utiliza para realizar el chocolate ecuatoriano es el mejor, que tenemos el mejor cacao del mundo, único en el mundo el mejor chocolate es de nuestro ecuador, el cacao de nuestras tierras le brinda ese sabor único, es trabajo de ecuatoriano para el mundo*. Dans cette question, l’*aficionado* aborde aussi les concepts de santé, de produit bio et de responsabilité sociale : *es sano consumirlo con frecuencia incluso si se tiene alguna enfermedad, es muy rico en su componente por ser*

*orgánico, cultivos orgánicos, calidad superior en contenido de antioxidantes y contenido nutricional, responsabilidad social, es bueno para la salud, amigable con el ambiente, que su cultivo no es un monocultivo por lo que su consumo promueve el comercio justo, exótico y delicioso y orgánico, es de buena calidad y tiene muchos nutrientes para la alimentación, es puro y sano para consumir, es un producto muy sano, le recomendaría un chocolate orgánico.* Il fait aussi de nombreuses références à la qualité du cacao *fino de aroma*. Les réponses de cette question sont aussi construites autour d'expressions plus démonstratives que descriptives démontrant ainsi le degré d'implication que l'*aficionado* a avec un produit qui fait partie de son patrimoine culturel et national.

La rédaction de la neuvième question (*Si debe elegir tres palabras o expresiones para definir el chocolate ecuatoriano, ¿Cuáles serían?*) oriente la réponse vers un choix plus terminologique que discursif incluant quelques expressions : *diversidad de sabores, sabor natural orgánico, consumir chocolate del ecuador, el mejor chocolate de los ecuatoriano, el mejor fino de aroma, alto sabor bajo precio alto contenido nutricional, el chocolate nos da energía, el mejor del mundo, algo único, con nuestra identidad, provoca felicidad por su componentes químicos, entre más amargo más saludable, el milagro del cacao, que nuestro producto tiene aroma de buena calidad, de exportación, tiene un aroma inigualable, tiene un sabor único, amor único por la delicia de este producto pasional, muy delicioso al paladar*. Dans ces expressions se retrouve le concept d'un produit organique équatorien sain inégalable qui procure du bonheur. Les termes significatifs les plus fréquents sont : *delicioso* (28), *rico* (25), *calidad* (20), *aroma* (20), *sabor* (20), *dulce* (19), *mejor* (15), *sabroso* (15), *único* (13), *exquisito* (11), *puro* (10), *agradable* (9), *aromático* (9), *saludable* (7), *ecuatoriano* (5), *fino* (5). Cette liste met en évidence que l'*Aficionado* utilise des termes communs dans un contexte spécifique pour partager ce que le mot *chocolate* évoque chez lui, aucun des termes de la liste n'appartient à aucune des roues des arômes (*Figure 13*).

Ces interprétation trouvent une résonance particulière dans les Hapax correspondants à ce sous-corpus sans stimulus. L'originalité des termes de fréquence d'utilisation égale

à un – les hapax – est qu'ils peuvent être considérés comme spécifiques au discours, car communément peu répandus. En effet, l'interprétation des Hapax peut varier selon la situation de communication, ils peuvent être considérés comme des mots-clés, des marqueurs de registre ou stylistiques. Ou encore mettre en évidence la présence d'un concept dans le corpus comme l'utilisation de diminutifs euphémiques pour limiter l'impact d'un commentaire négatif sur un produit équatorien : *amarguito, toquesito*. Ou même révéler des néologismes au service de l'imaginaire des *aficionados* comme *engordante, retrogusto*. Cependant l'évocation cognitive suscitée par le mot chocolat reste, parmi les hapax, fidèle à l'analyse des mots-clés du sous-corpus étudié. Il est possible d'identifier les références aux lieux de production : *amazonía, ecuador, pastaza, provincia, selvático, serrano* ; les références aux gens et aux activités sociales : *afrodisíaco, agriculteur, cariño, cumpleaños, enamoramiento, humano, mujer, ocasión, oficina, pareja, reunión, romanticismo, romántico, santo, sensualidad, social, ternura, valentin* ; les références à la santé : *ansiedad, antioxidante, apetitivo, bienestar, cuerpo, natural, salud, sano*. Il est remarquable de constater qu'il existe plus de termes liés au discours sensoriel dans les hapax du sous-corpus *aficionado* sans stimulus que dans le propre discours des producteurs : *alucinante, auténtico, deleitar, emoción, esparcimiento, exquisito* (erreur orthographique), *exquisitez, extraordinario, fabuloso, fantástico, fresco, gourmet, gustoso, inconfundible, inexplicable, infaltable, inigualable, insuperable, intenso, milagro, pasional, placentero, privilegio, rústico, semiamargo, superior, sutil*. De l'ordre du superlatif, les termes convoqués pour le discours sensoriel par les *aficionados* ne semblent pas avoir un format adéquat pour être organisés dans une roue des arômes ou des saveurs. Cependant, chaque discours a sa raison d'être, la terminologie qui permet la caractérisation précise d'un produit lors d'une session d'évaluation prend tout son intérêt pour situer un produit sur le marché. L'erreur consiste à vouloir utiliser ce discours lors des sessions de dégustation pendant lesquelles le système de valeur de l'*aficionado* n'est pas la description détaillée du produit, mais l'amplitude du plaisir ou de l'émotion ressentie. Pour une communication plus effective entre ces deux acteurs du domaine, le producteur de chocolat et

l'*aficionado*, la complémentarité entre un discours descriptif et un discours sensoriel serait une stratégie intéressante à considérer.

<b>Hapax - questionnaire sans stimulus</b>				
accesible	cumpleaños	gourmet	ocasión	reunión
afecto	deleitar	gustoso	oficina	romanticismo
afrodisíaco	detallista	helado	oloroso	romántico
agricultor	distinto	humano	organolépticas	rústico
almendra	ecuador	inconfundible	oscuro	salud
alucinante	emoción	inexplicable	pareja	sano
amable	enamoramiento	infaltable	pasional	santo
amarguito	engordante	inigualable	pastaza	selvático
amazonia	esparcimiento	insuperable	paz	semiamargo
ansiedad	espesor	intenso	perfecto	sensualidad
antioxidante	esquisito	intermediario	placentero	serrano
apetitivo	estilo	internacional	positivo	social
artesanal	exquisitez	lengua	postre	superior
auténtico	extraordinario	lleno	preparación	sutil
bienestar	fabuloso	mantecoso	primero	terapéutico
bombones	fantástico	menta	privilegio	ternura
bonito	felicidad	milagro	proveniente	tono
café	fragancia	mío	provincia	toquesito
cariño	fresco	monocultivo	pueblo	valentin
cosecha	fruto	mujer	relleno	viscosidad
cuerpo	fuerza	natural	retrogusto	vividios

Tableau 29 : Hapax de la réponse à la question treize du questionnaire (TXM)  
Source : élaboré par nous

L'analyse de sentiment obtenue avec le programme Orange (Figure 52) confirme le caractère fortement positif d'un passionné qui loue les particularités de son produit adoré. Avec une moyenne proche de 7, cette analyse démontre que le discours de l'*aficionado* est en harmonie avec la représentation sociale qu'il a du chocolat : l'orgueil, la joie et le plaisir. Si l'intention discursive du professionnel producteur de chocolat est de faire la promotion et de convaincre l'*aficionado* d'acheter son produit, il pourrait essayer d'en parler avec la même passion, s'approchant au mieux de son positivisme et de son enthousiasme.

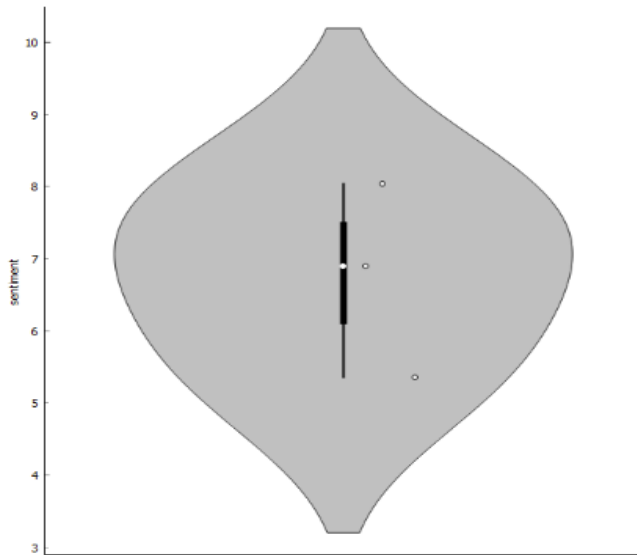


Figure 52 : Représentation de l'analyse de sentiment des textes du sous-corpus Aficionado sans stimulus (Orange)  
Source : élaboré par nous

#### 4.3. Réponses au questionnaire avec stimulus

Après avoir répondu à la partie cognitive du questionnaire, le participant retourne la feuille et l'intervieweur propose aléatoirement un échantillon de morceaux de chocolat identifiés comme *muestra 1* et *muestra 2* de façon à ne pas influencer le participant en lui indiquant le pourcentage de concentration en cacao du chocolat dégusté. L'échantillon numéro un a une concentration de 70% de cacao et le numéro deux de 100%. La dixième question propose une échelle de valorisation sur cinq de façon à pondérer les réponses des trois dernières questions (*Annexe 3 : Questionnaire avec stimulus*). Selon les valeurs obtenues, la barre de chocolat à 70% est évaluée à 4,12 sur 5 ; et la barre de chocolat à 100% est évaluée à 3,275 sur 5. Malgré l'amertume de cette dernière, le public montre une acceptation supérieure à la moyenne aux fortes concentrations en cacao ; résultat qui pourrait être différent dans un pays qui n'est pas producteur de cacao. Connaître les préférences du consommateur interviewé est une information primordiale au moment d'analyser le discours produit pour ne pas attribuer son appréciation comme qualité intrinsèque de la barre de chocolat, mais à la sensation

perçue au moment de la dégustation du produit étant conscient que les réponses à ce questionnaire ne peuvent être représentatives que des participants évoluant dans un écosystème linguistique situé. Cette information devient primordiale pour le producteur de chocolat quand il veut construire un discours promotionnel ciblé : les stratégies discursives optimales pour une barre à 70% ne le seront pas obligatoirement pour une barre à 100%. De plus, un consommateur ne parlera pas d'une barre de chocolat à 70% de la même façon qu'une barre à 100%, car la variation d'un des paramètres principaux de l'écosystème linguistique influence la production discursive comme le démontre l'analyse à suivre du sous-corpus *aficionado* avec stimulus.

Avec une valeur de 0,363, légèrement supérieure à la densité lexicale du sous-corpus représentatif du discours de l'*aficionado* sans stimulus (0,352), la densité lexicale du sous-corpus représentatif du discours de l'*aficionado* avec stimulus (*tableau 30*) reste comparable à la densité lexicale (0,358) du professionnel du producteur de chocolat (*tableau 10* et *tableau 26*).

sous-corpus SCQ avec stimulus	Token	Vocabulaire	Densité lexicale sur 1000 tokens aléatoires
Nombre total	3085	929	0.363

*Tableau 30 : Paramètres du sous-corpus de l'Aficionado avec stimulus*

*Source : élaboré par nous*

La moyenne de densité lexicale des sous-corpus avec et sans stimulus de l'*aficionado* (0,3575) est encore plus proche de celle du producteur de chocolat, démontrant ainsi que les deux protagonistes ont une variété lexicale sensiblement égale. Il reste à déterminer si leurs choix lexicaux sont tout aussi proches quand l'*aficionado* exprime une sensation.

En contact avec le produit, l'*aficionado* ne fait que peu de références au processus de fabrication : *proceso, fermentación, artesanal*, ce qui est en adéquation avec son discours issu de l'évocation (*Tableau 31*). Tout en dégustant le carré de chocolat,

l'*aficionado* n'hésite pas à se resservir un morceau pour définir les caractéristiques physiques que son palais (*paladar*), sa bouche (*boca*), ou sa langue (*lengua*) peuvent évaluer : *sabor, aroma, textura, olor, amargor, pureza, taza, característica, porcentaje, concentración, consistencia*. Il associe ces caractéristiques à des adjectifs communs afin de les évaluer : *sabor a chocolate auténtico, chocolate no muy dulce que en lo personal me resulta agradable y no empalaga, un sabor agradable sabor intenso, olor a cacao natural, su aroma a cacao y sabroso siento un aroma rico y sabor delicioso, se puede sentir su textura suave, cremosa, con su dulzura característica y su amargura al ser un chocolate puro, un sabor entre dulce y amargo, es un chocolate fino con características únicas, está muy delicioso, consistente y tiene agradable aroma, chocolate rico y puro delicioso aromático amargo maravilloso insaciable*. Cependant, mis à part la référence à l'odeur de café : *un olor a café, me hace sentir pasión tranquilidad al café*, l'*aficionado* n'utilise pas une terminologie spécifique pour identifier les arômes, les saveurs ou les odeurs et ne les décrit pas, comme le ferait un dégustateur professionnel, à l'aide d'une des nombreuses roues des arômes. Sa stratégie discursive se centre plus sur l'évocation plaisante que l'expérience sensitive. En effet, pour répondre à la question : *Probando el chocolate, ¿A qué le hace pensar?*, l'*aficionado* ne cherche pas une référence pour décrire la sensation olfactive ou gustative, mais communique ce que produit cette sensation en faisant appel à des souvenirs des plus agréables : *a una reunión familiar con una taza de chocolate lo mejor del mundo, felicidad aroma satisfacción de una buena taza de chocolate con leche, me recuerdo al chocolate que preparaba mi madre, me recuerda mi niñez, pensando para disfrutarlo con la familia y los amigos, la alegría de comer golosinas, mi esposa mis hijos cena en familia, su aroma penetra y se dirige al cerebro activando una sensación única de satisfacción, me hace pensar en mi niñez, infancia me hace pensar a lo rico que es el cacao como fruta al chocolate hecho en casa, excitante me hace pensar lo delicioso que es puesto que me hace recordar los premios que recibíamos cuando éramos niños, momentos felices al chocolate caliente que prepara mi mamá cuando hace frío, me hace sentir pasión tranquilidad, me recuerda a los chocolates que me solía comprar mi abuelita en la infancia un sabor agradable que incita a seguir comiendo, me da energía alegría, antes*



*del dulce el producto causa un estímulo de alegría y relajación, este chocolate tiene 2 matices el dulce de la vida con un toque de amargura.* Pendant la dégustation, les participants furent nombreux à fermer les yeux illustrant parfaitement par leur langage corporel ce commentaire écrit : *relajamiento rico que es algo relajante emite un olor que te transporta.* Cette stratégie discursive est aussi mise en évidence par la fréquence des termes : *alegría, bienestar, familia, infancia, niñez, satisfacción, casa, energía, feliz,* qui se trouvent être aussi présents dans le corpus sans stimulus démontrant ainsi qu'aussi bien l'évocation que la sensation réveille chez l'*aficionado* des souvenirs de bonheur partagés en famille.

Comme précédemment soulignés, les adjectifs choisis ne peuvent être définis seulement comme spécifiques à la terminologie du chocolat. Ces termes appartiennent au langage général, mais ils sont utilisés dans un écosystème linguistique spécifique lié au caractère sensoriel de la dégustation du chocolat. Même s'ils ne peuvent être catégorisés comme appartenant à une terminologie propre, ces adjectifs présentent des traits principaux qui permettent de catégoriser le type de qualification qu'ils apportent au substantif. Vassiliadou et Lammert (2011 : 8) proposent une description des informations sémantiques apportées par les adjectifs au substantif « odeur » : hédonique, effet, propriété, source, temporalité, intensité et identification. Ils précisent que le trait – propriété – recouvre plusieurs dimensions comme : propriété pure du substantif, sensorielle ou multisensorielle.

Proche de l'étude présentée sur les odeurs, les expériences sensorielles décrites par les *aficionados* font appel aux adjectifs hédoniques exprimant un jugement qualitatif principalement positif (tableau 31) : *bueno, rico, agradable, delicioso, fuerte, único, excelente, impresionante, incomparable, perfecto* ; et quand ils présentent une connotation négative, mis à part le terme *agrijo*, ils sont nuancés : *diferente, especial.*

[espos="NC"]: eslemma / F							
chocolate	109	pureza	6	familia	3	esencia	2
sabor	74	boca	5	infancia	3	experiencia	2
cacao	23	gusto	5	momento	3	fermentación	2

aroma	20	proceso	5	mundo	3	final	2
producto	19	taza	5	niñez	3	identidad	2
dulce	16	consumo	4	nivel	3	inicio	2
sensación	13	golosina	4	porcentaje	3	lengua	2
azúcar	12	mezcla	4	satisfacción	3	mañana	2
calidad	10	tiempo	4	agua	2	materia	2
leche	10	vez	4	bebida	2	pena	2
paladar	10	vida	4	capacidad	2	planta	2
textura	8	alegría	3	casa	2	repostería	2
olor	7	bienestar	3	concentración	2	tableta	2
amargor	6	café	3	consistencia	2	toque	2
persona	6	característica	3	energía	2	uso	2
<b>[espos="ADJ"]: eslemma / F</b>							
amargo	43	puro	11	cremoso	4	feliz	2
bueno mejor	33	sabroso	9	fino	4	gran grande	2
rico	22	excelente	7	orgánico	4	impresionante	2
agradable	21	agrio	6	químico	3	incomparable	2
delicioso	19	artesanal	6	suave	3	normal	2
dulce	16	diferente	6	amarguito	2	nuestro	2
fuerte	13	intenso	6	aromático	2	perfecto	2
natural	12	empalagoso	5	común	2	procesado	2
único	12	saludable	5	energético	2		
ecuatoriano	11	sano	5	especial	2		
<b>[espos="V.*"]: eslemma / F</b>							
ser	100	faltar	6	preparar	4	encantar	2
estar	22	haber	6	sentir	4	existir	2
hacer	19	poder	6	amargar	3	moderar	2
gustar	15	comer	5	conocer	3	parecer	2
tener	14	concentrar	5	estimular	3	pegar	2
probar	13	consumir	5	recordar	3	querer	2
sentir	13	decir	5	acompañar	2	saber	2
pensar	9	ir	5	activar	2	satisfacer	2
recomendar	7	quedar	5	apoyar	2	seguir	2
consumir	6	dar	4	comprar	2	superar	2
deber	6	mejorar	4	derretir	2	venir	2
<b>[espos="ADV"]: eslemma / F</b>							
muy	50	no	33	más	16	poco	7
bien	6	ya	6	antes	3	después	2
normalmente	2	también	2				

Tableau 31 : Lemmes de fréquence supérieure à 1 du sous-corpus *Aficionado* avec stimulus (TXM)

Source : élaboré par nous

Les adjectifs qualifiant une propriété présentent aussi une valence positive : *amargo*<sup>58</sup>, *dulce*, *natural*, *puro*, *(no) empalagoso*<sup>59</sup>, *sano*, *cremoso*, *fino*, *(sin) químico*<sup>60</sup>, *suave*, *aromático*. Les adjectifs identifiant la source mettent l'accent sur l'identité du produit (*nuestra golosina con sabor a Ecuador*) et son caractère naturellement bio. Le trait intensité : *fuerte*, *intenso*, *gran*, et le trait effet : *saludable*, *energético*, *feliz*, présentent un chocolat dont la concentration est bonne pour la santé, qui réveille et rend heureux.

Les verbes du discours avec stimulus présentent une plus grande variété, car *l'aficionado* décrit l'effet que procure la dégustation : *me recuerdo...*, *me hace pensar...*, *que preparaba mi...*, *parece un...*, *para disfrutarlo con...*, *a degustarlo*, *se derrite en la boca*, *me estimula*, *perfecto para acompañar...*, *me recuerda a los chocolates que me solía comprar mi*. Mis à part les adverbes intensificateurs qui sont aussi présents dans le corpus sans stimulus de façon à renforcer le sens et l'émphasis dans le discours, l'adverbe de négation fait son apparition pour soit changer la valence d'un adjectif : *no es empalagoso*, soit exprimer une opinion sur la forte concentration en cacao. En effet le chocolat 100% cacao présente une amertume qui n'est pas pour tous les palais : *no fue lo que pensé*, *no me gustó muy amargo*, *no me gustó para nada*. Dans ce cas, *l'aficionado* met en évidence une brèche entre sa représentation d'un chocolat parfait et la sensation procurée par le 100% cacao, le verdict est sans appel même si certains essaient de minimiser l'effet étant donné qu'ils savent que c'est un produit équatorien : *no será del gusto de cualquiera*.

---

<sup>58</sup> L'amertume est une qualité positive pour les amateurs de chocolat qui apprécient une forte concentration en cacao qui doit être amer sans être âcre (*agrio*).

<sup>59</sup> Le terme *empalagoso* (écœurant) possède une forte valence négative, mais dans le corpus il est associé à la négation *no*.

<sup>60</sup> Le terme *químico* possède une valence négative quand il est associé à un produit alimentaire. Cependant, dans le corpus étudié, il est associé à la préposition privative *sin* donnant ainsi une valence positive à l'expression *sin químico*.



Comme premier mot-clé, l'adjectif hédonique *rico* exprime un jugement qualitatif positif du produit qui n'est plus seulement évoqué, mais senti (*tableau 32*). Il manifeste une première réaction qui, dans la même catégorie hédonique, est suivie de *agradable, delicioso, sabroso, excelente*. Les substantifs *calidad, pureza, aroma, textura, porcentaje, taza y olor* apparaissent comme des caractéristiques propres du produit qui semblent expliquer la sensation perçue aussi décrite comme une expérience. Ces propriétés sont renforcées par les adjectifs *natural, fuerte, puro, amargo, ecuatoriano, cremoso, concentrado, intenso, único, artesanal, (no) empalagoso*, qui complètent la description du produit. Les autres substantifs présents comme mots-clés font référence à l'endroit où naît la sensation : *boca, paladar* ; et aux ingrédients : *cacao, azúcar, chocolate y leche*. Les verbes – *siente, merecer (la pena), impregna (que se impregna en toda la boca), gustó, probar* – de cette liste sont étroitement liés à la sensation, mais sont aussi présents les verbes – *pensar, diría* – qui introduisent la description de l'effet en accord avec *experiencia* et *niñez*.

Word	YAKE!	Word	YAKE!	Word	YAKE!	Word	YAKE!
rico	0,07	cremoso	0,03	concentrado	0,03	artesanal	0,02
calidad	0,06	cantidad	0,03	experiencia	0,03	empalagoso	0,02
pureza	0,06	final	0,03	intenso	0,03	pensar	0,02
sensación	0,06	impregna	0,03	chocolate	0,03	persona	0,02
cacao	0,06	textura	0,03	gusta	0,03	probarlo	0,02
natural	0,05	delicioso	0,03	boca	0,03	única	0,02
fuerte	0,05	punto	0,03	paladar	0,03	porcentaje	0,02
puro	0,05	sabrosa	0,03	pega	0,03	sabroso	0,02
agrio	0,05	gustó	0,03	niñez	0,03	taza	0,02
agradable	0,05	siento	0,03	bebida	0,03	excelente	0,01
siente	0,05	pastoco <sup>61</sup>	0,03	recomendado	0,03	olor	0,01
amargo	0,04	mejores	0,03	capacidad	0,03	leche	0,01
azúcar	0,04	merece	0,03	nivel	0,03	dulce	0,01
ecuatoriano	0,04	pena	0,03	sabor	0,02	diría	0,01
aroma	0,04	presentaciones	0,03	único	0,02		

Tableau 32 : 59 mots-clés du sous-corpus *Aficionado* avec stimulus (*Orange*)

Source : élaboré par nous

<sup>61</sup> Synonyme de *masilla* qui se traduit en français par : mastic ; *pastoso* peut être associé à patafix.

Une fois la sensation décrite, l'*aficionado* revient à l'expérience vécue, pas à partir d'une évocation, mais à partir du ressenti. Présente dans le discours avec ou sans stimulus, l'expérience, le plus souvent agréable, semble être ce que l'*aficionado* recherche dans la consommation du produit. En conclusion, le discours du professionnel gagnerait en pertinence, en relation à la troisième maxime conversationnelle de Grice (1979 : 57-72), s'il se référait à l'expérience que procure le produit.

Le regroupement hiérarchique (*figure 54*) de ces mots-clés confirme l'analyse précédente. Le regroupement C1 réunit les verbes qui décrivent aussi bien la sensation que son effet : *sentir, pegar, impregnar gustar*. Le regroupement C2 met en évidence les caractéristiques qui permettent de sentir et vivre l'expérience que procure la dégustation d'un produit unique, pur, naturel et équatorien (regroupement C3 : *único, puro, natural, ecuatoriano*) qui en bouche devient crémeux, excellent, bon, savoureux, intense, et fort (regroupement C5 : *cremoso, excelente, rico, sabroso, intenso, fuerte*). Le regroupement C4 fait référence à la boisson chocolatée qui reste une des façons traditionnelles que les Équatoriens ont de consommer ce produit : *bebida, taza, leche, chocolate, cacao, azúcar*.

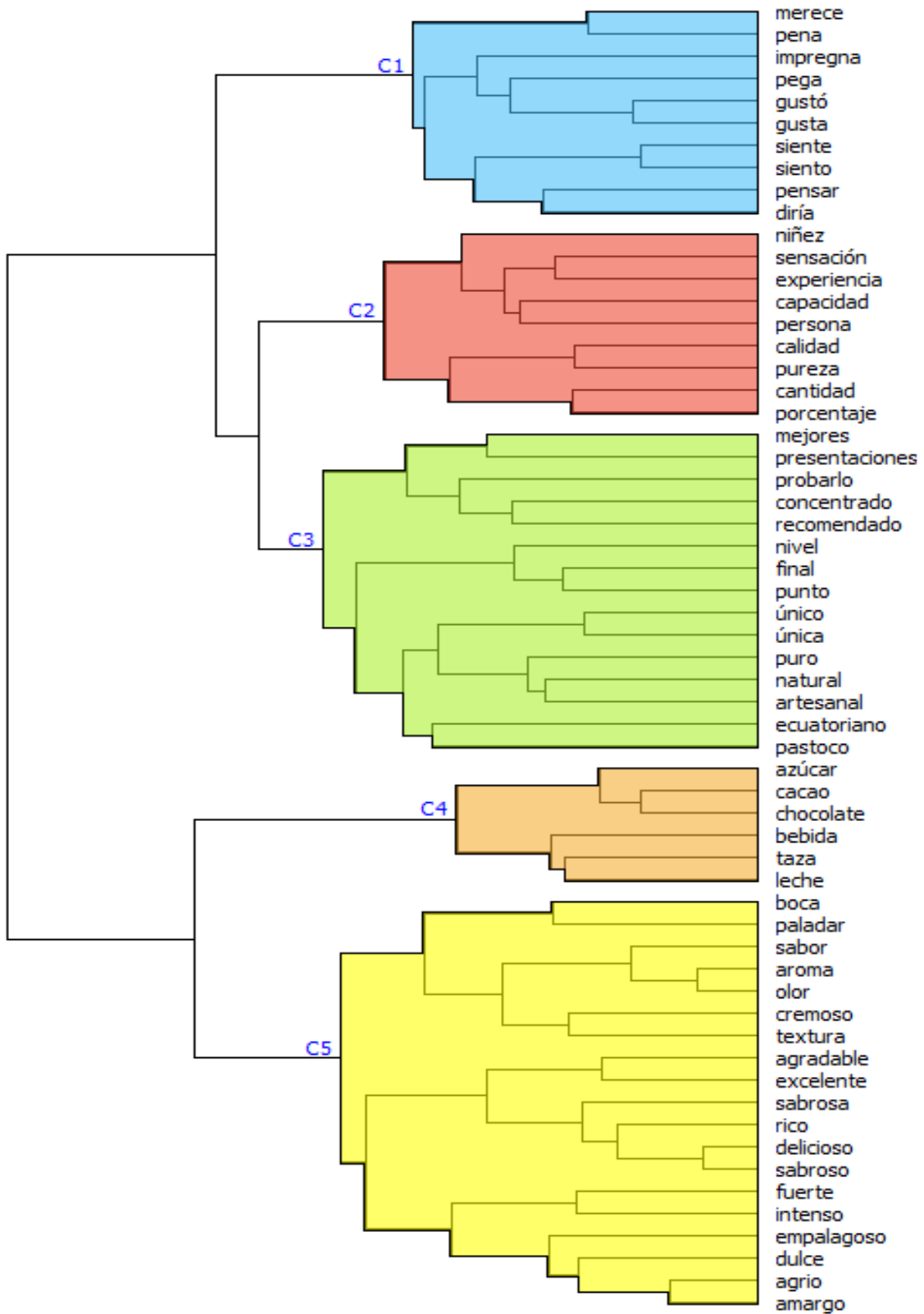
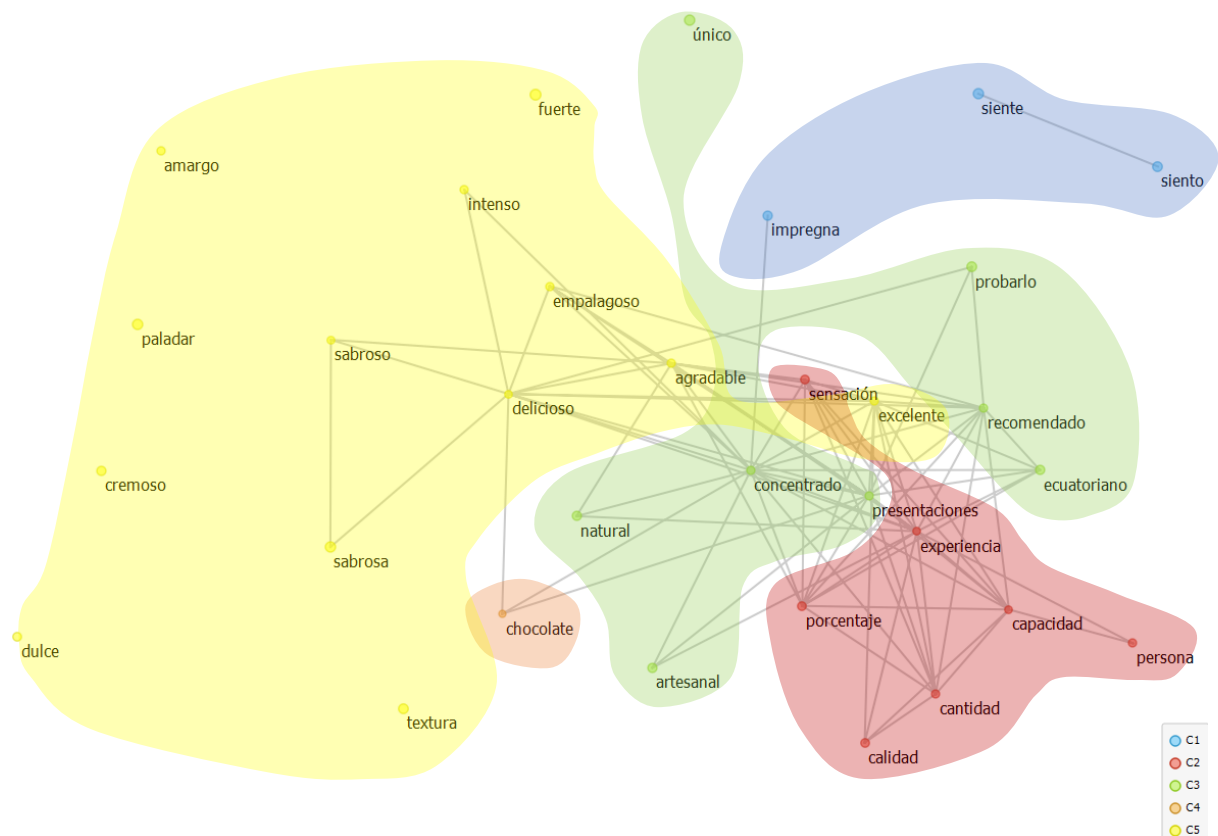


Figure 54 : Regroupement hiérarchique des mots-clés du sous-corpus aficionado avec stimulus (Orange)

Source : élaboré par nous

La fonction d'échelle multidimensionnelle (Borg et coll., 2018), *multidimensional scaling (MDS)* met en évidence le lien central entre C2, C3 et C5 (*Figure 55*) renforçant l'idée que les caractéristiques aussi bien physiques que sensorielles offrent une expérience placentaire décrite par les adjectifs hédoniques. L'*aficionado* n'analyse pas le goût ou l'odeur à l'aide d'une roue des arômes, il ressent et partage son expérience en utilisant des adjectifs à forte valence positive avant de décrire son plaisir par analogie à une situation de bonheur qui fait partie de son expérience culturelle personnelle. Par conséquent, si le professionnel producteur de chocolat veut s'adresser aux *aficionados*, ses clients potentiels, il gagnera en pertinence, s'il s'exprime dans leur écosystème linguistique pour les convaincre que le produit leur fera vivre une expérience inoubliable tout en utilisant leurs adjectifs et expressions hédoniques.



*Figure 55 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus Aficionado avec stimulus (Orange)*  
*Source : élaboré par nous*



La terminologie spécialisée des roues des arômes restera d'une grande utilité pendant les concours et les dégustations entre experts pour déterminer des paramètres de comparaison et d'évaluation entre différents produits. Cependant, elle risque d'être incomprise aussi bien par les consommateurs que par les producteurs de cacao.

Comme pour le sous-corpus *aficionado* sans stimulus, le sous-corpus *aficionado* avec stimulus présente une dimension limitée pour l'analyse des cooccurrences des mots-clés. Cependant, l'outil *Network explorer* du programme Orange offre une représentation graphique des termes dont la fréquence de cooccurrence est supérieure à 5 dans une fenêtre de quatre mots à gauche et à droite, considérant les termes dont la fréquence est supérieure à 10 (*figure 56*).

Selon ce réseau sémantique, pour les *aficionados* participant au questionnaire, le chocolat équatorien *fino de aroma* présente une riche (bonne), agréable et délicieuse saveur amère et sucrée ; ses caractéristiques en font un produit sain et unique. Dans cette représentation graphique du réseau des cooccurrences, les expressions qui se réfèrent aux expériences ne peuvent être représentées, car étant propres aux différents participants, elles ne présentent pas des fréquences suffisantes. Cependant, dans les réponses de la question onze : *Probando el chocolate, ¿A qué le hace pensar?*, il est possible de lire (*Annexe 18* : sous-corpus questionnaire avec stimulus) : *una reunión familiar con una taza de chocolate lo mejor del mundo, satisfacción de una buena taza de chocolate con leche, me recuerdo al chocolate que preparaba mi madre, golosina con sabor a ecuador, me recuerda mi niñez, pensando para disfrutarlo con la familia y los amigos, que es algo relajante, las cosas buenas de la vida, mi niñez la alegría de comer golosinas, mi niñez la alegría de comer golosinas, en realidad no pienso en nada solo siento el sabor del chocolate, mi esposa mis hijos cena en familia, quien fue el inventor de tan deliciosa golosina, me hace pensar en mi niñez, infancia, me hace recordar los premios que recibíamos cuando éramos niños, momentos felices al chocolate caliente que prepara mi mamá cuando hace frío, me hace sentir pasión tranquilidad, me recuerda a los chocolates que me solía comprar mi abuelita en la infancia.*

Ces exemples mettent en évidence que même après avoir déguster le chocolat, l'*Aficionado* continue de mettre en œuvre une stratégie discursive autour d'expressions évocatrices d'actuations sociales riches en émotions qui font appel à la propre expérience de celui qui pose la question. Le participant continue de chercher une résonance chez celui qui pose la question pour partager ses émotions plus que pour les décrire.

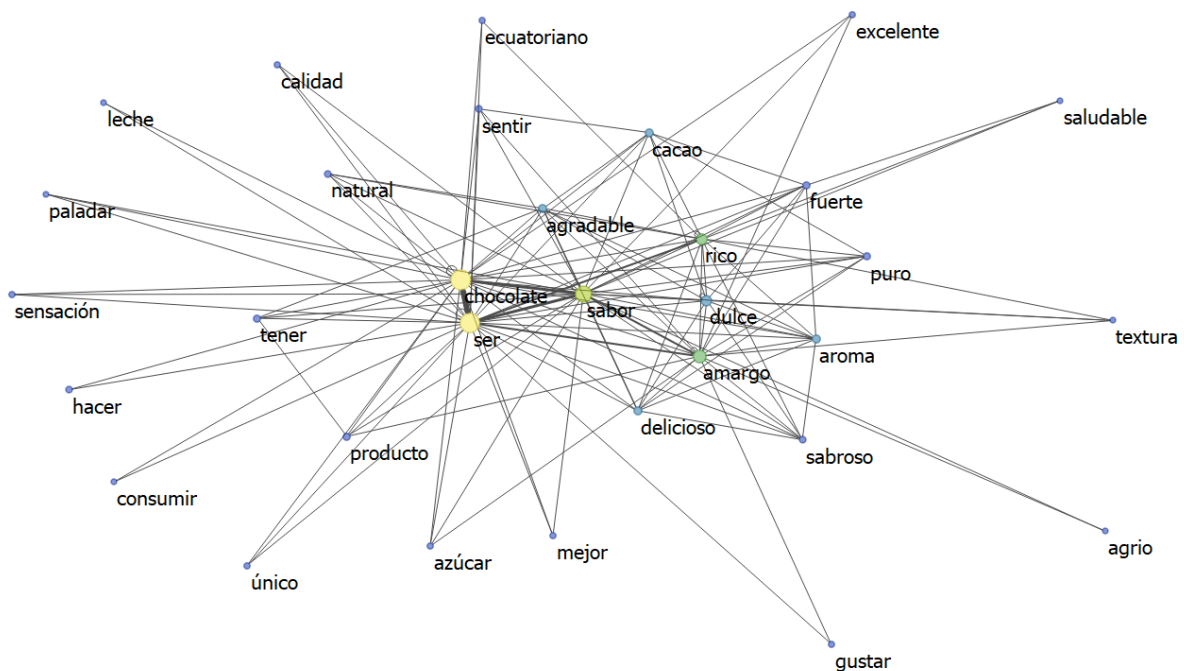


Figure 56 : Réseau sémantique des principales cooccurrences du sous-corpus Aficionado avec stimulus (Orange)

Source : élaboré par nous

Dans l'ensemble des réponses à la douzième question : *si debe explicar a un amigo lo que siente, ¿qué diría?*, il est possible de lire (Annexe 18 : sous-corpus questionnaire avec stimulus) : *prueba y veras que te va a encantar, se siente un sabor que perdura, que nuestro chocolate es uno de los mejores del mundo, que no entiendo a las personas cuando dicen que no les gusta, el producto causa un estímulo de alegría y relajación, te activa las papillas gustativas, es necesario para sentirse bien y animado chocolate*

*muy bueno, que este chocolate tiene dos matices el dulce de la vida con un toque de amargura, emoción una experiencia única está bueno, ...*; de la même manière, dans les réponses présentées se retrouvent de nombreuses références à la fierté nationale de produire un des meilleurs cacao au monde; à ses qualités physiques, naturelles et sensorielles. Les exemples relevés mettent en évidence l'importance de l'effet du sensoriel sur le bien-être de celui qui le déguste. Les mots-clés des réponses à la treizième question confirment les analyses précédentes. Cette liste de mots (*tableau 33*) classés selon leurs fréquences d'usage dans les réponses à la treizième question du questionnaire appliqué ; représente la base terminologique utilisée par les aficionados pour définir le chocolat qu'ils ont goûté. Une fois encore, elle démontre que la spécificité terminologique du discours ne réside pas dans le terme, mais dans l'écosystème linguistique dans lequel il est utilisé.

Word	Freq.	Word	Freq.	Word	Freq.
amargo	31	gustar	6	amor	3
rico	30	único	5	tranquilidad	2
dulce	20	textura	5	sano	2
bueno	20	suave	5	relajante	2
sabor	31	cremoso	5	recomiendo	2
delicioso	16	calidad	5	producto	2
aroma	16	concentrado	4	proceso	2
chocolate	13	agrio	4	placer	2
sabroso	12	sensación	3	paladar	2
excelente	11	químico	3	identidad	2
agradable	11	orgánico	3	fragrante	2
fuerte	10	intenso	3	fino	2
saludable	8	feo	3	experiencia	2
puro	8	falta	3	ecuatoriano	2
natural	8	energético	3	casero	2
exquisito	8	delicia	3	artesanal	2
aromático	7	cacao	3		

*Tableau 33 : 50 premiers mots de la réponse à la question treize du questionnaire (TXM)  
Source : élaboré par nous*

Il est aussi remarquable dans le sous-corpus avec stimulus que cette réalité soit aussi mise en évidence par les termes de fréquence d'utilisation égale à un qui indiquent la rareté d'un événement comme la présence de certains commentaires négatifs (*tableau 34*) : *aceitoso, insípido, desabrido, desagradable, pastoso, empalagoso*. Ou encore mettre en évidence la présence du concept des diminutifs euphémiques qui limitent l'impact d'un commentaire négatif sur un produit équatorien : *poquito, amarguito*. Ils peuvent être propre à l'écosystème linguistique sans pour autant qu'ils soient considérés comme des termes de spécialité : *cuencano, nacional, waorani*. Cependant, comme les hapax compilés (*tableau 34*) le démontrent, la majorité met en valeur le processus de recherche cognitive que certains *aficionados* mettent en œuvre pour exprimer soit leurs émotions ressenties : *perdurable, consistente, equilibrada, maravilloso, estimulo, profundo, agridulce, excitante, exótico, refrescante, delicado, satisfactorio, fresco, diferente, genial, oloroso, increíble, insaciable, perceptible*, soit leurs émotions évoquées : *encanta, energía, adictivo, esposa, afrodisíaco, pureza, familiar, salud, felicidad, anti-estrés, auténtico, café, golosina, casa, dulzura, pasional, gratificante*.

Loin d'être représentatifs d'une terminologie spécialisée décontextualisée, les hapax sont ici aussi considérés comme représentatifs d'une terminologie générale située dans un discours sensoriel spécialisé construit autour du chocolat en contexte équatorien.

<b>Hapax - questionnaire avec stimulus</b>				
aceitoso	comprar	encanta	insípido	perdurable
adecuada	concentración	energía	ligeramente	perfecto
adictivo	consistente	equilibrada	mantiene	poquito
aditivos	consumo	esposa	maravilloso	probar
afrodisíaco	convierte	estimulo	materia	profundo
agridulce	cristalización	excitante	mejor	pureza
alimenticio	cuencano	exótico	mejorar	refrescante
amarguito	dark	familiar	moldeable	salud
anímico	delicado	felicidad	nacional	satisfactorio
anti-estrés	desabrido	fermentación	nuevo	superado
auténtico	desagradable	fresco	oscuro	tendencia
azúcar	diferente	genial	olor	trabajo
café	diluir	golosina	oloroso	tristeza

casa	dulzura	gran	pasional	vez
cobre	duradero	gratificante	pastoso	volveré
color	duro	increíble	paz	waorani
comestible	empalagoso	insaciable	perceptible	

Tableau 34 : Hapax de la réponse à la question treize du questionnaire (TXM)

Source : élaboré par nous

L'analyse de sentiment obtenue avec le programme Orange (*Figure 57* : Représentation de l'analyse de sentiment des textes du sous-corpus Aficionado avec stimulus (Orange)) confirme qu'après la dégustation du produit, *l'aficionado* reste très positif dans son évaluation. Cependant, même si en moyenne l'analyse de sentiment des sous-corpus avec et sans stimulus reste équivalente, l'analyse du sous-corpus avec stimulus présente une plus grande amplitude. Le fait de proposer un échantillon de chocolat dont la concentration en cacao est de 100% peut expliquer cette différence indiquant ainsi que son amertume ne correspond pas toujours à la représentation sociale que le terme chocolat a déclenché dans la partie cognitive du questionnaire.

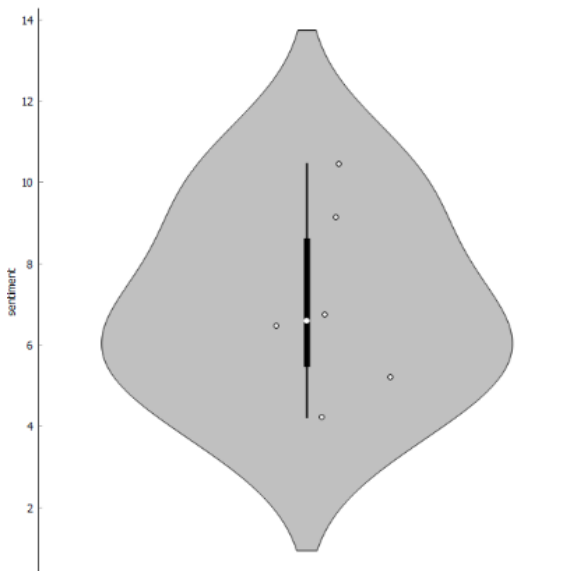


Figure 57 : Représentation de l'analyse de sentiment des textes du sous-corpus Aficionado avec stimulus (Orange)

Source : élaboré par nous

L'amertume d'un chocolat noir 100% cacao peut surprendre et ne pas être en phase avec la représentation évoquée par le simple fait d'entendre le mot chocolat. Cette hypothèse est aussi soutenue par des adjectifs présents dans le discours comme *feo* et par l'évaluation indiquée par la question neuf du questionnaire (*tableau 25*). Cependant, dans son ensemble, cette analyse démontre que le discours de l'aficionado reste en harmonie avec la représentation sociale qu'il a du chocolat.

## Conclusion

---

La thématique du moule discursif du producteur de chocolat dans un écosystème linguistique situé au salon du chocolat à Quito est essentiellement liée au processus de production, aux ingrédients utilisés ainsi qu'à leurs caractéristiques. Dans le sous-corpus compilé, l'analyse textométrique ne met pas en évidence l'utilisation d'une terminologie spécifique, mais d'une terminologie commune utilisée dans un écosystème linguistique spécifique dans lequel sont construits les réseaux sémantiques de chaque terme choisi. Il a été observé que la stratégie de communication du discours sensoriel reste limitée et se restreint à l'utilisation de verbes sensoriels (goûter, tester, se démarquer, sentir, adorer, rappeler, évoquer, ressembler). Les seuls termes dont la fréquence d'utilisation est significative qui se réfèrent à une sensation olfactive et gustative sont *floral* et *frutal*. Les noms propres présents dans ce discours font référence à des lieux de production, donnant ainsi une situation géographique précise à l'écosystème linguistique ; et à des marques de producteur de chocolat qui sont aujourd'hui reconnues aussi bien en Équateur qu'à l'extérieur soulignant ainsi la renommée de cet écosystème linguistique dans le marché international du chocolat. La stratégie discursive du producteur de chocolat s'inscrit dans la promotion du produit, son intention de communication est de convaincre l'audience de consommer et d'acheter. Sûr des qualités de son produit, le producteur de chocolat invite à goûter ce qui lui semble un argument suffisant étant donné qu'il n'éprouve pas le besoin de développer son discours au-delà de la dégustation. Il est ici important de se rappeler que les entrevues ont été réalisées pendant le salon du chocolat à Quito sur les stands des professionnels, il aurait été surprenant que cette invitation à la dégustation n'apparaisse pas comme stratégie de communication. La garantie de la qualité de leurs produits repose sur l'origine et le type de cacao présenté comme une référence internationale : *le cacao nacional fino de aroma*. Le processus de fabrication présente aussi une certaine importance, s'opposant au concept d'industrialisation massive, le producteur met

l'accent sur l'artisanat et la durabilité du processus utilisé en adéquation avec l'Amazonie et son système ancestral de culture : la *chakra*. Il existe une différence d'utilisation entre le pluriel et le singulier des termes arôme et saveur : le singulier est majoritairement associé au cacao et le pluriel au chocolat. Cette nuance suggère que l'apport du processus de transformation du cacao en chocolat ne se limite pas à une simple plus-value économique, cela sous-entend que le savoir-faire du producteur permet de révéler tout le potentiel du cacao, de multiplier les arômes et les saveurs. Cependant, fort d'une longue tradition de production du cacao, dans un domaine où la production de chocolat en Équateur peut être considérée comme récente, la stratégie discursive du producteur de chocolat reste centrée sur les caractéristiques de la fève. L'analyse met en évidence que le discours sensoriel est particulièrement absent du discours du producteur de chocolat. Bien que positif, le discours présente le produit comme le résultat d'un processus de transformation de qualité, d'une fève de cacao de qualité, sans décrire l'effet sensoriel et l'impact émotionnel que le produit peut susciter chez le consommateur.

La première considération à prendre en compte pour comparer le discours du producteur de chocolat avec le discours du producteur de cacao est la différence essentielle entre les écosystèmes linguistiques où sont produits les échanges : le salon du chocolat à Quito versus la plantation de cacao dans les *chakra* des communautés amazoniennes. Même si ces deux écosystèmes linguistiques se situent sur le sol équatorien, ils n'ont que peu de points communs : le thème, la langue pour les entrevues réalisées en espagnol, et l'interlocuteur qui conduit l'entrevue ; les autres paramètres sont diamétralement opposés. Comme premier résultat de ces différences, il est possible d'observer dans le discours des producteurs de cacao des emprunts au kichwa qui est la langue maternelle des interlocuteurs, l'utilisation de suffixes et une prononciation marquée par le lieu géographique et par le kichwa. Il n'est pas surprenant de constater que les producteurs de cacao ne font pas de distinction dans le choix du pluriel ou du singulier pour le terme saveur étant donné que peu d'entre eux ont eu l'opportunité de pouvoir déguster le chocolat produit à partir de la fève de cacao qu'ils vendent. Avant



d'être producteurs de cacao, ils sont agriculteurs, ils appliquent un modèle de pluri-culture adapté aux conditions climatiques et à la biodiversité amazonienne expliquant ainsi la fréquence d'utilisation de substantifs d'autres produits dans leur discours. La présence d'une nature exubérante dans l'écosystème linguistique du producteur de cacao se traduit par l'utilisation d'adjectif comme naturel ou organique pour décrire la saveur de la fève produite. L'Amazonie est aussi connue pour sa richesse en métaux précieux, dès l'enfance les membres des communautés apprennent à laver le sable des rivières pour chercher de l'or. L'analogie de la fève de cacao avec la pépite d'or, directement liée au rendement économique, est ainsi naturelle dans cet environnement. Sans une compréhension détaillée de leur écosystème linguistique, il est difficile d'interpréter le discours du producteur de cacao. Sa réalité sociale ainsi que l'organisation de la chaîne de production du chocolat permettent d'expliquer l'absence d'un discours sensoriel dans ses réponses. Comme agriculteur communautaire, il se réfère au processus de production de la fève, à son intérêt économique, il éprouve des difficultés à s'exprimer sur le chocolat qui ne fait pas partie de sa réalité. Même si son discours reste positif, il est nuancé par les conditions de son quotidien.

La comparaison entre les discours des producteurs de cacao et de chocolat reflète la sectorisation de la chaîne de fabrication du chocolat en Équateur, la réalité socioéconomique de ses acteurs et la relation qu'ils entretiennent avec le produit fini. Cependant, malgré leur différence, aucun des deux acteurs n'inclut dans sa production langagière un discours sensoriel spécifique. Il est important de prendre en considération que traditionnellement, l'industrie chocolatière exportait le cacao et actuellement elle commence à le transformer, ce qui peut être un argument pour justifier que le discours soit centré sur les particularités de la fève. Cependant, le discours des producteurs de chocolat gagnerait en pertinence s'il s'appuyait sur les stratégies discursives exprimées par leurs clients potentiels : les *aficionados*.

En effet, en harmonie avec la représentation sociale qu'ils ont du chocolat : l'orgueil, la joie et le plaisir, la stratégie discursive de ces derniers reste plus centrée sur l'évocation d'une expérience placentaire de sensations, de plaisir et de satisfactions que sur

l'expérience sensitive du moment de dégustation. Sans réellement inclure dans son discours une stratégie terminologique spécifique du cacao ou du chocolat, les *aficionados* mettent en œuvre des expressions évocatrices d'actes sociaux qui sont des sources d'émotions particulières dont l'intensité ne peut être perçue dans son intégralité que par eux-mêmes. Cependant, dans leur discours ils cherchent une résonance avec leur interlocuteur, celui qui leur pose la question, de façon à partager cette émotion plus qu'à l'expliquer. Il est ici important de souligner que l'interlocuteur qui réalise l'entrevue est un compatriote qui partage le même écosystème linguistique permettant ainsi que cette recherche de résonance semble naturelle et légitime. Il serait intéressant de réaliser ce questionnaire avec un interlocuteur étranger à l'écosystème linguistique des *aficionados* pour évaluer l'impact de cette variation sur leur stratégie discursive. Le professionnel centre son discours sur le processus et le produit, l'*aficionado* centre son discours sur sa propre expérience et la représentation sociale qu'il a du produit. L'intention de communication du professionnel est de convaincre, l'intention de communication de l'*aficionado* est de partager. En conclusion, pour convaincre, la stratégie discursive du professionnel gagnerait en pertinence si sa première intention de communication était fondée sur l'expérience placentaire de sensations, de plaisir et de satisfactions qu'il partagerait comme *aficionado*.

De la réflexion ontologique et épistémologique initiale à l'interprétation croisée des résultats, le processus cognitif inscrit dans la linguistique sensorielle présente des contributions d'aspect théorique, méthodologique et pragmatique.

Une des interrogations légitimes que le chercheur peut avoir à l'esprit tout au long de l'étude se centre sur la place du spécialisé dans la production discursive de locuteurs qui s'expriment dans un contexte professionnel autour d'un thème spécifique comme le cacao ou le chocolat. Cette réflexion, d'aspect théorique, déjà soulevée par Petit (2010), interroge directement la nature du discours spécialisé. Dans le domaine professionnel, est-il possible de réduire le spécialisé d'un discours à la simple spécialisation de sa terminologie ? Selon notre position, une des contributions d'aspect théorique de cette étude est d'offrir une réponse négative à la question précédente : le discours spécialisé

n'est pas que déterminé par les définitions terminologiques présentes dans les ouvrages de référence. Selon les résultats obtenus, la valeur sémantique d'un terme dépend de la spécialisation et la spécificité de l'écosystème linguistique dans lequel il est utilisé, des stratégies de production langagière, des intentions de communication, et de la valeur sémantique que les termes choisis renvoient aux interlocuteurs de cet écosystème. Dans l'étude présentée, il a été démontré que le syntagme *pepa de oro* n'aura pas la même épaisseur sémantique (Bach, 2021a : 63), pour un producteur, un professionnel ou un *aficionado* selon sa représentation sociale, la situation, la stratégie et l'intention de communication. Même si une personne a répondu que l'expression pépite d'or était utilisée au niveau national (voir la contribution C-03 du paragraphe 1.2.1 Représentation sociale du cacao et du chocolat en Équateur), le producteur de la fève de cacao ne connaît pas vraiment la valeur réelle de ce qu'il cultive étant donné qu'il n'a pas eu le plaisir de déguster le chocolat fabriqué à partir de sa pépite d'or (voir la contribution C-05). L'agriculteur va donc utiliser cette expression comme il l'a reçu, c'est-à-dire comme un patrimoine national qui ne s'inscrit pas dans son écosystème sociolinguistique. Contrairement au producteur de cacao, le producteur de chocolat s'approprie la représentation sociale de l'expression « pépite d'or » comme stratégie d'exportation de son produit fini. Dans son écosystème sociolinguistique, cette expression s'inscrit dans la terminologie spécifique du domaine du chocolat pour identifier les particularités du cacao *Arriba fino de aroma* : pour le producteur de chocolat équatorien l'épaisseur sémantique de *pepa de oro* est comparable à l'épaisseur sémantique que la notion de « climat » peut avoir pour les producteurs de vin bourguignon.

Il peut être extrapolé à partir de cette réflexion que plus un terme peut être utilisé dans de nombreux écosystèmes linguistiques sans grandes variations de sa valeur sémantique, plus il sera « commun » ou moins « spécifiques », plus un terme présente un nombre limité d'écosystèmes linguistiques dans lesquels il peut être utilisé sans changement drastique de sa valeur sémantique, et plus il sera confiné à un domaine spécifique. La spécificité terminologique ne réside pas dans la définition du terme, mais

dans l'écosystème linguistique dans lequel il est employé, par qui il est employé, comment et pourquoi il est employé. Cette réponse s'inscrit l'approche d'une linguistique cognitive située développée dans les différents travaux d'analyse du « spécialisé » de l'équipe « Modèles et Discours » du laboratoire dijonnais TIL – Centre Interlangue – Texte, Image, Langage (EA 4182) de l'Université de Bourgogne<sup>62</sup>.

L'obtention des données s'inscrit dans un cadre méthodologique triphasé dans lequel la conceptualisation du thème étudié prend forme à travers l'étude d'un « corpus ambiant » construit à partir des discours produits sur internet. Cette première phase offre au chercheur une description plus ou moins spécifique, selon les paramètres de sélection des sites internet considérés, de l'empreinte discursive générale que la mise en discours d'un concept utilisé laisse dans l'écosystème linguistique du domaine de spécialité étudié. La deuxième phase consiste à compiler un corpus de textes dans lesquels le concept étudié est cognitivement évoqué dans différentes situations et intentions de communication qui correspondent à une variété précise d'écosystèmes linguistiques représentatifs du domaine de spécialité. Dans l'étude présente, cette phase est représentée par le discours des entrevues faites aux producteurs ainsi que par la première partie du questionnaire appliqué aux *aficionados*. La troisième phase naît d'un stimulus physique présenté comme le déclencheur de la production d'un discours spécifique réalisé par un locuteur spécialisé dans le domaine. Dans le cas de cette étude, cette phase est représentée par le discours de l'*aficionado* produit après avoir dégusté un morceau de chocolat. De cognitif, le stimulus devient sensitif, cette distinction permet la comparaison entre la production langagière déclenchée par un processus cognitif avec celle déclenchée par un processus sensitif donnant ainsi accès à la mise en discours des émotions créées à partir d'évocations et de sensations. Cette méthodologie triphasée a permis de présenter les réseaux sémantiques de l'expression de la sensorialité dans le discours spécialisé construit autour du chocolat en contexte équatorien en partant d'un sous-corpus ambiant général vers un sous-corpus représentatif des différentes situations spécifiques particulières à la chaîne de production du chocolat en Équateur. De cette

---

<sup>62</sup> <https://til.u-bourgogne.fr/thematiques/axe-de-recherche-3.html>

manière, l'analyse donne au chercheur un accès à l'épaisseur sémantique du discours sensoriel par la superposition de l'étude des différentes strates de sa réalité écolinguistique. Selon Hernández Sampieri et Mendoza (2018), cette méthodologie s'inscrit dans un type de recherche comparative descriptive dans lequel l'obtention de l'information ne se base pas sur la manipulation de variable, mais sur la compilation, l'observation et l'interprétation d'une base de données pour décrire les caractéristiques d'un phénomène. Appliquée à la linguistique de corpus, la compilation permet l'obtention d'un corpus de textes représentatifs d'un écosystème linguistique. La textométrie, et ses outils informatiques d'aide à l'analyse de discours, permet une observation quantitative du corpus compilé offrant ainsi au chercheur des résultats dont l'interprétation qualitative s'inscrit parfaitement dans le cadre théorique d'une linguistique cognitive sensorielle située. La complémentarité des cadres théoriques de la linguistique de corpus, de la textométrie et de la linguistique cognitive sensorielle située au service d'un processus d'analyse mixte aussi bien quantitatif que qualitatif est une contribution importante pour le cadre méthodologique d'une linguistique du discours sensoriel qui cherche à interpréter dans l'empreinte discursive les stratégies d'expression du ressenti que sa source soit une évocation ou une sensation.

La contribution d'aspect pragmatique est directement liée à la pertinence des résultats pour les professionnels du secteur, et plus spécifiquement, pour les producteurs de chocolat dont l'intention finale de communication est de convaincre un potentiel acheteur d'acquiescer son produit. Dans cette étude, l'interprétation des résultats permet de conclure que la connaissance des stratégies de production langagière permet au professionnel d'orienter son discours promotionnel pour qu'il soit congruent avec celui de son client potentiel. En d'autres termes, que le discours du producteur de chocolat gagnerait en pertinence s'il incluait dans sa production langagière les stratégies discursives de l'*aficionado* tout en s'inspirant de son intention de communication inscrite dans le partage d'expériences placentaires, de joie et de satisfaction. En valorisant ce résultat, le producteur de chocolat applique une sectorisation de son discours selon l'intention et la situation de communication. En effet, en présence

d'experts-dégustateurs, le langage technique de descripteurs caractéristiques des propriétés physiques, gustatives et olfactives du chocolat dans un contexte d'évaluation et de comparaison de produits prend toute son importance, car il sert l'intention de communication qui est le positionnement du produit sur un marché international hautement compétitif. Ce langage technique de descripteurs, souvent représenté par une roue des arômes ou de saveurs, devient impropre si l'intention de communication est de convaincre un potentiel client, qui ne partage pas l'écosystème linguistique de l'expert-dégustateur, d'acheter le produit. Autre grande différence entre le langage technique et le langage sensoriel réside dans la généralisation et la spécificité. Pour comparer des produits dans un espace économique international, il semble légitime de normaliser la technicité des termes allant jusqu'à la circonscrire dans une langue franche – l'anglais – pour être intégrée dans un espace économique international, comme le montrent les étiquettes présentées dans le paragraphe 1.1.3. L'écologie du langage comme rapport entre langue et humain/société. En raison de son caractère international, ce discours technique perd toute relation culturelle et devient obsolète dans une situation de communication où le professionnel producteur de chocolat doit s'adresser à un public particulier qui évolue dans un écosystème linguistique fortement marqué par sa culture et ses traditions. Dans ce cas, le discours produit doit être situé et doit résonner avec l'interlocuteur tout comme le discours de l'*aficionado* résonne avec ceux qui partagent son écosystème linguistique.

Le discours est une entité sémantique qui nous permet d'organiser « cognitivement le monde en lui donnant un sens » (Adam, 2014 : 296), il devient un témoin de l'existence sociale du locuteur, « il permet une confrontation des visions cognitives (pensées, autres cultures présentes et, plus largement, autres personnes). Les textes sont des lieux de mémorisation. D'archivage, de conservation et de transmission de savoir » (Adam, 2014 : 296). Conformément à sa nature, le discours se présente comme un accès privilégié à l'intention de communication de l'acteur social et à l'interprétation qu'il possède de sa réalité à travers le prisme de l'écologie linguistique dans laquelle il évolue. Dans le cadre de la linguistique cognitive sensorielle, chaque analyse de

discours compilé en corpus permet de saisir et interpréter les stratégies discursives qui maintiennent une cohérence et une cohésion entre sensations, émotions, et stratégies d'expressions sensorielles à partir de l'environnement physique et socioculturel qui détermine l'écologie du langage. Au-delà du cacao, du chocolat, et même du vin souvent pris comme référence, l'application du processus de recherche utilisé dans cette étude à d'autres écosystèmes linguistiques liés à différents produits qui évoquent ou stimulent des émotions partagées dans un discours sensoriel comme le café, l'huile d'olive, le thé pour les saveurs, allant jusqu'à inclure les épices et les parfums pour les odeurs, pourrait confirmer, compléter ou améliorer les contributions d'aspect théorique, méthodologique et pragmatique présentées et ouvrir de nouvelles voies de recherche dans le domaine de l'expression du sensoriel située dans un discours spécialisé.

# Table des matières

Introduction.....	1
Première Partie : Cadre théorique et méthodologique pour la caractérisation des discours sensoriels autour du chocolat en contexte équatorien .....	11
Chapitre 1 : Écologie de la langue située dans la filière du chocolat en Équateur .....	13
1.1 Le paradigme de l'« Écologie du langage » .....	14
1.1.1 Définition, évolution et positionnement épistémologique de l'écologie du langage. ....	14
1.1.2. L'écologie du langage comme rapport entre langue et environnement naturel .....	30
1.1.3. L'écologie du langage comme rapport entre langue et humains/société ....	35
1.1.4 Discussion de la pertinence de ces approches pour le travail de thèse, de la construction et de l'analyse du corpus .....	40
1.2 L'approche ethnographique de la spécialité.....	45
1.2.1. Représentation sociale du cacao et du chocolat en Équateur.....	47
1.2.2. Description du cacaoyer et de son fruit.....	50
1.2.3. Histoire du cacao en Équateur .....	53
1.2.4. Chocolat actuellement produit en Équateur.....	57
1.3 Processus de transformation artisanale et industrielle .....	61
1.3.1. Processus de transformation artisanal du cacao.....	63
1.3.2. Processus de transformation industriel du cacao .....	65
Chapitre 2 : Le corpus en linguistique sensorielle.....	72
2.1 Positionnement théorique de la recherche : Linguistique sensorielle et écologie du langage .....	74



2.2 Typologie de la terminologie sensorielle selon trois grands ensembles .....	84
2.3 Processus de compilation du corpus d'étude .....	87
Deuxième Partie : Caractérisation analytique du corpus et modélisation .....	107
Chapitre 3 : Analyse textométrique et interprétation croisée des corpus considérés .	109
3.1 Moule discursif représentatif du sous-corpus discours <i>ambient</i> .....	113
3.2 Dualisme dans le corpus professionnel selon l'activité, l'intention de communication et la langue utilisée.....	142
3.2.1. Moule discursif représentatif du discours professionnel du producteur de chocolat .....	143
3.2.2. Moule discursif représentatif du discours professionnel du producteur de cacao en espagnol.....	161
3.2.3. Moule discursif représentatif du discours professionnel du producteur de cacao en kichwa .....	177
Chapitre 4 : Empreinte du discours sensoriel dans les situations communicatives situées de l' <i>aficionado</i> .....	191
4.1. Stratégies discursives de l' <i>aficionado</i> .....	191
4.2. Réponses au questionnaire sans stimulus .....	192
4.3. Réponses au questionnaire avec stimulus .....	205
Conclusion .....	222

## Liste des tableaux

Tableau 1 : Paramètres des situations de communication sélectionnées pour compiler le corpus d'étude.....	93
Tableau 2 : Paramètres du sous-corpus questionnaire .....	98
Tableau 3 : Paramètres du sous-corpus professionnel .....	101
Tableau 4 : Paramètres de transcription .....	102
Tableau 5 : Paramètres du sous-corpus ambiant.....	114
Tableau 6 : 100 premiers lemmes des noms communs, adjectifs, verbes et adverbess du sous-corpus ambiant (TXM).....	117
Tableau 7 : 100 premiers mots-clés du sous-corpus ambiant (Orange).....	120
Tableau 8 : cooccurrence de cacao dans le sous-corpus ambiant (TXM) .....	136
Tableau 9 : cooccurrence de chocolat dans le sous-corpus ambiant (TXM) .....	138
Tableau 10 : Paramètres du sous-corpus professionnel producteur de chocolat .....	144
Tableau 11 : Lemmes du sous-corpus professionnel producteur de chocolat dont la fréquence est supérieure à 10 (TXM).....	146
Tableau 12 : 100 premiers mots-clés du sous-corpus professionnel producteur de chocolat (Orange) .....	149
Tableau 13 : cooccurrence de cacao dans le sous-corpus professionnel producteur de chocolat (TXM).....	159
Tableau 14 : cooccurrence de chocolat dans le sous-corpus professionnel producteur de chocolat (TXM).....	159
Tableau 15 : Paramètres du sous-corpus professionnel producteur de cacao.....	162
Tableau 16 : Lemmes du sous-corpus professionnel producteur de cacao dont la fréquence est supérieure à 10 (TXM).....	164
Tableau 17 : 100 premiers mots-clés du sous-corpus professionnel producteur de cacao en espagnol (Orange).....	167
Tableau 18 : cooccurrence de cacao dans le sous-corpus professionnel producteur de cacao (TXM).....	173
Tableau 19 : concordances de fino dans le sous-corpus professionnel producteur de cacao (TXM).....	174

Tableau 20 : concordances de « elaborado a partir de su cacao? » dans le sous-corpus professionnel producteur de cacao (TXM).....	175
Tableau 21 : cooccurrence de chocolate dans le sous-corpus professionnel producteur de cacao (TXM).....	175
Tableau 22 : Paramètres du sous-corpus professionnel producteur de cacao.....	177
Tableau 23 : Comparaison de la production discursive en espagnol et en kichwa.....	178
Tableau 24 : Mots les plus utilisés du sous-corpus professionnel producteur de cacao en kichwa (Orange) .....	184
Tableau 25 : Profil de l'aficionado .....	192
Tableau 26 : Paramètres du sous-corpus de l'Aficionado sans stimulus.....	192
Tableau 27 : Lemmes de fréquence supérieure à 1 du sous-corpus Aficionado sans stimulus (TXM) .....	194
Tableau 28 : 41 mots-clés du sous-corpus Aficionado sans stimulus (Orange) .....	197
Tableau 29 : Hapax de la réponse à la question treize du questionnaire (TXM).....	204
Tableau 30 : Paramètres du sous-corpus de l'Aficionado avec stimulus .....	206
Tableau 31 : Lemmes de fréquence supérieure à 1 du sous-corpus Aficionado avec stimulus (TXM) .....	210
Tableau 32 : 59 mots-clés du sous-corpus Aficionado avec stimulus (Orange).....	212
Tableau 33 : 50 premiers mots de la réponse à la question treize du questionnaire (TXM).....	218
Tableau 34 : Hapax de la réponse à la question treize du questionnaire (TXM).....	220

## Liste des figures

Figure 1 : Territoires des peuples et nationalités autochtones, afro-équatoriens et montubios.....	28
Figure 2 : Zone de production du cacao .....	33
Figure 3 : géographie de l'Équateur .....	34
Figure 4 : Étiquette de chocolat vendu en Équateur .....	38
Figure 5 : Étiquette chocolat « au lait » vendu en Équateur.....	39
Figure 6 : Production mondiale de fèves de cacao en volume de 2016/2017 à 2021/2022 .....	57
Figure 7 : Processus de collecte des fèves .....	61
Figure 8 : Processus de fermentation.....	62
Figure 9 : Processus de séchage.....	62
Figure 10 : Processus artisanal : de la fève à la tablette de chocolat.....	65
Figure 11 : Exemple de séchoir industriel .....	66
Figure 12 : Processus industriel de la fève à la tablette de chocolat .....	68
Figure 13 : Roues des arômes du chocolat, de l'huile d'olive et du vin.....	78
Figure 14 : Écologie linguistique de la production langagière en linguistique sensorielle .....	83
Figure 15 : Paramètre du tertium comparationis .....	90
Figure 16 : Mots-clés sélectionnés pour constituer le corpus ambiant.....	94
Figure 17 : Association de mots-clés sélectionnés pour faire une recherche sur Internet .....	94
Figure 18 : Paramètre de la recherche sur Internet pour compiler le corpus ambiant..	95
Figure 19 : Exemples d'URL sélectionnés dans Bootcat pour compiler le corpus ambiant .....	95
Figure 20 : Flux d'analyse des sous-corpus (Orange) .....	112
Figure 21 : Nuage des cent premiers lemmes du sous-corpus ambiant (Orange) .....	118
Figure 22 : Flux de travail pour le Nuage de mots du sous-corpus ambiant (Orange) .....	119

Figure 23 : Flux de travail appliqué aux 100 premiers mots et aux mots-clés du sous-corpus ambiant (Orange) .....	122
Figure 24 : Exemple de vectorisation des 20 premiers mots les plus fréquents (Orange) .....	123
Figure 25 : Matrice de distance des 20 premiers mots les plus fréquents (Orange)...	124
Figure 26 : Regroupement hiérarchique des 100 premiers mots du sous-corpus ambiant (Orange).....	127
Figure 27 : Regroupement hiérarchique des mots-clés du sous-corpus ambiant (Orange) .....	131
Figure 28 : Agrandissement de la projection bidimensionnelle (MDS) des 100 premiers mots du sous-corpus ambiant (Orange) .....	133
Figure 29 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus ambiant (Orange).....	134
Figure 30 : Réseau sémantique des principales cooccurrences du sous-corpus ambiant (Orange).....	139
Figure 31 : Représentation de l'analyse de sentiment des textes du sous-corpus ambiant (Orange).....	142
Figure 32 : Nuage des cent premiers lemmes du sous-corpus professionnel producteur de chocolat (Orange).....	148
Figure 33 : Regroupement hiérarchique des 100 premiers mots du sous-corpus professionnel producteur de chocolat (Orange).....	152
Figure 34 : Regroupement hiérarchique des mots-clés du sous-corpus professionnel producteur de chocolat (Orange) .....	154
Figure 35 : Agrandissement de la projection bidimensionnelle (MDS) des mots du sous-corpus producteur de chocolat dont la fréquence est supérieure à dix (Orange).....	156
Figure 36 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus producteur de chocolat (Orange) .....	158
Figure 37 : Réseau sémantique des principales cooccurrences du sous-corpus professionnel producteur de chocolat (Orange).....	160
Figure 38 : Représentation de l'analyse de sentiment des textes du sous-corpus professionnel producteur de chocolat (Orange).....	161

Figure 39 : Nuage des cent premiers lemmes du sous-corpus professionnel producteur de cacao en espagnol (Orange) .....	166
Figure 40 : Regroupement hiérarchique des 100 premiers mots du sous-corpus professionnel producteur de cacao (Orange) .....	171
Figure 41 : Agrandissement de la projection bidimensionnelle (MDS) des cent premiers mots du sous-corpus producteur de chocolat (Orange) .....	172
Figure 42 : Représentation de l'analyse de sentiment des textes du sous-corpus professionnel producteur de cacao (Orange) .....	176
Figure 43 : Nuage des cent premiers termes du sous-corpus professionnel producteur de cacao en kichwa (Orange).....	180
Figure 44 : Regroupement hiérarchique des mots fréquents du sous-corpus professionnel producteur de cacao (Orange).....	186
Figure 45 : Photo d'une boisson chocolatée avec son muchin prise à la communauté « Porvenir ».....	187
Figure 46 : Agrandissement de la projection bidimensionnelle (MDS) des mots les plus utilisés du sous-corpus producteur de cacao (Orange) .....	188
Figure 47 : Réseau sémantique des principales cooccurrences du sous-corpus professionnel producteur de cacao (Orange) .....	189
Figure 48 : Nuage des cent premiers lemmes du sous-corpus Aficionado sans stimulus (Orange).....	196
Figure 49 : Regroupement hiérarchique des mots-clés du sous-corpus Aficionado sin stimulus (Orange) .....	198
Figure 50 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus Aficionado sin stimulus (Orange).....	199
Figure 51 : Réseau sémantique des principales cooccurrences du sous-corpus Aficionado sans stimulus (Orange) .....	200
Figure 52 : Représentation de l'analyse de sentiment des textes du sous-corpus Aficionado sans stimulus (Orange) .....	205
Figure 53 : Nuage des cent premiers lemmes du sous-corpus Aficionado avec stimulus (Orange).....	211

Figure 54 : Regroupement hiérarchique des mots-clés du sous-corpus aficionado avec stimulus (Orange) .....	214
Figure 55 : Agrandissement de la projection bidimensionnelle (MDS) des mots-clés du sous-corpus Aficionado avec stimulus (Orange) .....	215
Figure 56 : Réseau sémantique des principales cooccurrences du sous-corpus Aficionado avec stimulus (Orange) .....	217
Figure 57 : Représentation de l'analyse de sentiment des textes du sous-corpus Aficionado avec stimulus (Orange) .....	220

## Bibliographie

Abad, A., Acuña, C., et Naranjo, E. (2020). El cacao en la Costa ecuatoriana: estudio de su dimensión cultural y económica. *Estudios de la Gestión: revista internacional de administración*, 7, 59-83. <https://doi.org/10.32719/25506641.2020.7.3>

Adamou, E. (2016). *A corpus-driven approach to language contact: Endangered languages in a comparative perspective* (Vol. 12). Walter de Gruyter GmbH et Co KG.

Aitchison J. (2012). *Words in the mind: An introduction to the mental lexicon*. Wiley-Blackwell, 352 p. ISBN: 978-1-118-17096-0

Alcívar Rodríguez, B.; Pino Peralta, S. et Aguilar Azuero, R. (2019). *El gran cacao de los pequeños productores. Una mirada al escenario agro socioeconómico de los cacaoteros en la Cordillera Oriental del Ecuador*. Guayaquil: Ediciones Holguín. ISBN 978-9942-14-510-9.

Altmann, P. (2012). Interculturalidad y plurinacionalidad como conceptos decoloniales-colonialidad y discurso del movimiento indígena en el Ecuador. In *XV Encuentro de Latinoamericanistas Españoles*, 131-138. Trama editorial; CEEIB. Madrid, España. halshs-00873607

Andy Alvarado, P., Calapucha Andy, C., Calapucha Cerda, L., López Shiguango, H., Tanguila Andy, A., et Tanguila Andy, D. (2012). *Sabiduría de la cultura kichwa de la amazonia ecuatoriana Tomo II*. Cuenca: Universidad de Cuenca.

Angué, K. (2009). Rôle et place de l'abduction dans la création de connaissances et dans la méthode scientifique peircienne. *Recherches qualitatives* 28(2), 65-94.

Bach, M. (2020a). Sensorial discourse and corpus in the digital humanities era: The example of the wine language. *Digital Scholarship in the Humanities*, 35 (1), 1-16. <https://doi.org/10.1093/llc/fqz006>.



Bach, M. (2020b). Les descripteurs sensoriels d'une langue à l'autre. Enjeux cognitifs pour la traduction. *Cognitivism et Traductologie Approches sémantiques et psychologiques*, 6, 59-76. Publisher : Classiques Garnier. DOI : 10.15122. ISBN : 978-2-406-09944-4

Bach, M. (2021a). *Vers une Sémantique Discursive cognitive. Réflexions théoriques et applications empiriques sur un corpus de langue allemande*. (Doctoral dissertation, University of Burgundy).

Bach, M. (2021b). Discours du sensoriel au prisme de l'approche cognitive. Tour d'horizon et perspectives. *Travaux neuchâtelois de linguistique*, 75, 105-124.

Baker, P., Egbert, J., (2016). *Triangulating Methodological Approaches in Corpus-Linguistic Research*. New York: Routledge.

Ballesteros, L. A. (2008). La ecolingüística: a propósito del portugués brasileño. *Actas de las I jornadas de ecología y lenguajes*. Córdoba, Universidad Nacional de Córdoba. ISBN 978-950-33-0736-6

Bastardas-Boada, A. (2005). *Cap a una sostenibilitat lingüística*. Barcelona: Editorial Angle.

Bastardas-Boada, A. (2014). Ecología y sostenibilidad lingüísticas: una aproximación desde la (socio)complejidad. Miguel Casas Gómez (dir.), Rocío Vela Sánchez (ed.) *Memoriam. XIV Jornadas de Lingüística*, Cádiz, Universidad de Cádiz, 143-163. ISBN 978-84-9828-458-4.

Bastardas-Boada, A. (2016). Ecología lingüística y lenguas minorizadas: algunas notas sobre el desarrollo del campo. Fernández Planas, A. Ma. (ed.) *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, 451-458. ISBN: 978-84-608-9830-6.

Bastardas-Boada, A. (2017). The ecology of language contact: minority and majority languages. Fill, A., Penz, H. (eds.) *The Routledge Handbook of Ecolinguistics*, New York: Routledge, 451-458. ISBN: 9781138920088.

Bejarano, D., Llanos, A., Rubio, R., et Bonilla, J. (2018). Protocolo de transcripción ortográfica CLICC. *Instituto Caro y Cuervo*, 1-35.

Biber, D. (2015). Corpus-based and corpus-driven analyses of language variation and use. Bernd Heine, and Heiko Narrog (eds), *The Oxford Handbook of Linguistic Analysis*, <https://doi.org/10.1093/oxfordhb/9780199677078.013.0008>

Biber, D., Conrad, S., et Reppen R. (1998). *Corpus Linguistics, Investigating Language Structure and Use*. Cambridge : Cambridge University Press. ISBN 0-521-49957-7.

Blanchet, P. (2014). Ce que le plurilinguisme des espaces francophones nous apprend sur l'enseignement du français : l'impérative ouverture à la pluralité contre les idéologies glottophobes. *Le français, une langue pour réussir*. Rennes : Presses universitaires de Rennes. ISBN : 9782753557383. DOI : <https://doi.org/10.4000/books.pur.65561>.

Boas, F. (1911). *Introduction [to Handbook of American Indian Languages]* (No. 677). US Government Printing Office.

Boas, F. (1938). *Handbook of American Indian languages* (No. 40). US Government Printing Office.

Bojanowski, P., Grave, E., Joulin, A., et Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5.

Borg, I., et Groenen, M. (2018). *Applied Multidimensional Scaling and Unfolding*. Cham: Springer International Publishing. ISSN 2191-5458. <https://doi.org/10.1007/978-3-319-73471-2>

Brémond, J. (2011). Vignobles et vins de Rioja : histoire(s), identité(s), image(s) et stratégie(s). Territoires du vin, 3. <http://preo.u-bourgogne.fr/territoiresduvin/index.php?id=1415>

Brown, G., Yule, G. (1983). *Discourse Analysis*. Cambridge : Cambridge University Press.

Caballero, F. S., et Maldonado, C. (Eds.). (2016). *Comunicación, decolonialidad y buen vivir* (Vol. 7). Ediciones Ciespal. ISBN 9789978551455

Cadena Fiallos, M. F. (2015). *Diseño del proceso para la elaboración de una bebida energética a partir de excedentes de cacao*. Trabajo final para la obtención del título: Ingeniera de Alimentos Espol Facultad de Ingeniería en Mecánica y Ciencias de la Producción, Guayaquil, 72 p.

Cahuana Velasteguí, A. L.; Cruz Fernández J.; Méric, O. et Gautier, L. (2021a). Mettre le chocolat en mots : terminologie de filière ou commerciale ? D. M. De Vecchi and V. Delavigne (eds.), *Entreprises et organisations : perspectives terminologiques*, 85-107. Paris : Presse Sorbonne Nouvelle. ISBN: 978-2-37906-067-0. <https://shs.hal.science/halshs-03249490/>

Calvet, L.-J. (1999). *Pour une écologie des langues du monde*. Paris, Plon, 304 p.

Calvet, L.-J. (2019). Louis-Jean Calvet. L'écologie des langues. Nicolas Journet éd., *Les Grands Penseurs du langage*, 121-124. Auxerre : Éditions Sciences Humaines. <https://doi.org/10.3917/sh.journ.2019.01.0121>

Campos, R., Mangaravite, V., Pasquali, A., Jatowt, A., Jorge, A., Nunes, C. et Jatowt, A. (2020). YAKE! Keyword Extraction from Single Documents using Multiple Local Features. *Information Sciences Journal*. Elsevier, 509, 257-289. Condamines, A., Narcy-Combes, J.-P. (2015). La linguistique appliquée comme science située. F. Carton, J.-P. Narcy-Combes, M.-F. Narcy-Combes, D. Toffoli (eds.), *Cultures de recherche en linguistique appliquée*, 209-229. Paris : Riveneuve éditions. ISBN: 978-2-36013-354-3.

Danesi, M. (2021). *Linguistic relativity today: Language, mind, society, and the foundations of linguistic anthropology*. Routledge

Digonnet, R. (2018). *Pour une linguistique sensorielle*. Paris : Honoré Champion. ISBN : 978-2-74534-729-9. <http://digital.casalini.it/9782745347299>.

Dubois, D. (2010). *Le sentir et le dire. Concepts et méthodes en psychologie et linguistique cognitives*. Paris: l'Harmattan, 391. ISBN: 978-2-29624-600-3.

Eisenstein, J. (2019). *Introduction to Natural Language Processing*. MIT Press.

Everaert-Desmedt, N. (2011). La sémiotique de Peirce dans Louis Hébert. *Signo (en ligne)*.  
<http://www.signosemio.com/peirce/semiotique.asp>

Fairclough, N. (2005). Discourse analysis. *organizational studies: the case for critical realism*, *Organization Studies*, 26(6), 915-939.

Fernández, A. G. (2018). Big data y corpus lingüísticos para el estudio de la densidad léxica. *SKOPOS Revista Internacional de Traducción e Interpretación*, 9, 107-122. E-ISSN: 2695-8465. ISSN: 2255-3703.

Fillmore, C. J. (1992). Corpus linguistics' vs. 'computer-aided armchair linguistics'. *Directions in Corpus Linguistics. Proceedings of Nobel Symposium*, 82, 35-60.

Fix, U. (1990). Die Kategorie Stil als theoretisches Problem : Zur Einführung. *Beiträge zur Stiltheorie*, Leipzig: Verlag Enzyklopädie, 7–18.

Fix, U. (1999). Textsorte – Textmuster – Textmuster Mischung. Konzept und Analysebeispiele. *Cahiers d'Études Germaniques*, 37(2), 11-26.

Fornel, M. D., et Léon, J. (2000). L'analyse de conversation, de l'ethnométhodologie à la linguistique interactionnelle. *Histoire épistémologie langage*, 22(1), 131-155

Gardner, H. (1987). *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Gautier, L. (2009). Nochmals zum (Fach-)Textmuster: von der Kognition zur Beschreibung einzelner Textexemplare. *Lylia*, numéro spécial Histoires de textes, 1–8.  
<http://langues.univ-lyon2.fr/1184-Histoires-textes.html>.

Gautier, L. (2018). Émotions, expressivité et évaluation. La triple face (cachée ?) des descripteurs sensoriels : l'exemple du discours de présentation/dégustation de vins. *Symposium Connaître vs Ressentir. Affects, émotions et expressivité en discours spécialisés*, Lyon: Université Lumière Lyon II. <https://halshs.archives-ouvertes.fr/halshs-01722225>.

Gautier, L., Bach, M. (2017). La terminologie du vin au prisme. Des corpus oraux de dégustation / présentation (français-allemand) : Entre émotions, culture et sensorialité. *Études de Linguistique Appliquée*, 188, 485-509.

Gautier, L., Le Fur, Y, et Robillard, B. (2015). La minéralité du vin : mots d'experts et de consommateurs. *Unité et diversité dans le discours sur le vin en Europe: Actes du colloque d'Innsbruck*, 210 p, Francfort/Main: Peter Lang, 149-168.

Gautier, L. (2019). La recherche en « langues-cultures-milieus » de spécialité au prisme de l'épaisseur socio-discursive. Calderón M., Konzett-Firth, C., *Dynamische Approximationen. Festschriftliches pünktlichst zu Eva Lavrics 62,5. Geburtstag*. ISBN : 978-3-631-79649-8.

Gledhill, C., Patin S., Zimina, M. (2017). Lexico-grammaire et textométrie : identification et visualisation de schémas lexico-grammaticaux caractéristiques dans deux corpus juridiques comparables en français. *Corpus*. <http://journals.openedition.org/corpus/2868>. DOI : <https://doi.org/10.4000/corpus.2868>.

Glynn, D., et Fischer, K. (Eds.). (2010). *Quantitative methods in cognitive semantics: Corpus-driven approaches* (Vol. 46). Walter de Gruyter.

Grave, E., Bojanowski, G., Joulin, A., et Mikolov, T. (2018). Learning Word Vectors for 157 Languages. *Proceedings of the International Conference on Language Resources and Evaluation*. <https://doi.org/10.48550/arXiv.1802.06893>.

Grice, H.P (1979). Logique et conversation. *Communications*, 30, 57-72.

Grzech, K. Z. (2016). *Discourse enclitics in Tena Kichwa: A corpus-based account of information structure and epistemic meaning* (Doctoral dissertation, SOAS, University of London).

Guilhaumou, J. (1996). Vers une histoire des événements linguistiques. Un nouveau protocole d'accord entre l'historien et le linguiste. *Histoire épistémologie langage*, 18(2), 103-126.

Guilhaumou, J. (2000). De l'histoire des concepts à l'histoire linguistique des usages conceptuels. *Genèses*, (1), 105-118.

Haarmann, H. (1990). Language planning in the light of a general theory of language: a methodological framework. *International Journal of the sociology of Language*, 1990(86), 103-126. <https://doi.org/10.1515/ijsl.1990.86.103>

Habert, B. (2000). Des corpus représentatifs : de quoi, pour quoi, comment ? *Linguistique sur corpus Études et réflexions*, Ed. M. Bilger, Perpignan : Presses Universitaires de Perpignan, 31, 11–58.

Hagège, C. (2006). *Combat pour le français : au nom de la diversité des langues et des cultures*. Editions Odile Jacob

Halliday, M. A. K. (2006). New ways of meaning: The challenge to applied linguistics. A. Fill, P. Mühlhäusler (eds.), *Ecolinguistics reader: Language, ecology, and environment*, 175-202. London and New York: Continuum.

Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146-162.

Haugen, E. (1972). *The ecology of language*. Stanford, Calif: Stanford University Press.

Heiden S. (2010). The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. *24th Pacific Asia Conference on Language, Information and Computation*. Sendai, Japan. [http://halshs.archives-ouvertes.fr/docs/00/54/97/64/PDF/paclic24\\_sheiden.pdf](http://halshs.archives-ouvertes.fr/docs/00/54/97/64/PDF/paclic24_sheiden.pdf).

Heiden S., Magué J.-P., et Pincemin B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. *JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data*, 12, Rome, Italie. [http://halshs.archives-ouvertes.fr/docs/00/54/97/79/PDF/Heiden\\_al\\_jadt2010.pdf](http://halshs.archives-ouvertes.fr/docs/00/54/97/79/PDF/Heiden_al_jadt2010.pdf).

Hernández Sampieri, R., et Mendoza, C. (2018). *Metodología de la investigación, las rutas cuantitativa cualitativa y mixta*. México, México: Mc Graw Hill. ISBN: 978-1-4562-6096-

Honnibal, M. (2013). *A good Part-of-Speech Tagger in about 200 lines of Python*.  
<https://explosion.ai/blog/part-of-speech-pos-tagger-in-python>

Hymes D. (1974). *Social anthropology, sociolinguistics and the ethnography of speaking*.  
Foundations in Sociolinguistics, an Ethnographic Approach, Philadelphia, University of  
Pennsylvania Press, 83-117.

Itier, C. (2017). *Diccionario quechua sureño castellano*. Lima: Editorial Commentarios.

Jean-Marie E., Bereau D, et Robinson J-C. (2021). Benefits of Polyphenols and  
Methylxanthines from Cocoa Beans on Dietary Metabolic Disorders. *Foods*, 10(9), 2049.  
<https://doi.org/10.3390/foods10092049>.

Jo, Y. H. (2013). La descolonialidad y el Sumak Kawsay en Ecuador. *Revista  
Iberoamericana*, 24(1), 37-63.

Junyent, M. C. (1991). *Les llengües del món: Ecolingüística*. Barcelona: Empúries. ISBN:  
978-8-475-96182-8.

Jurafsky, D., Martin, J.H. (2019). Vector semantics and embeddings. *Speech and Language  
Processing: An Introduction to Natural Language Processing, Computational Linguistics,  
and Speech Recognition*, 94-122.

Kleiber, G. (2012). Carte d'identité linguistique des odeurs. *IRIS*, 33, 91-103.

Kress, G. (1985). *Linguistic processes in sociocultural practice*. Deakin, Australia: Deakin  
University Press.

Kress, G. (1988). *Communication and culture: An introduction*. Kensington, Australia:  
University of New South Wales Press.

Krupnik, I. (2011). 'How many Eskimo words for ice?' Collecting Inuit sea ice  
terminologies in the International Polar Year 2007–2008. *The Canadian Geographer/Le  
Géographe canadien*, 55(1), 56-68.

Krupnik, I., Müller-Wille, L. (2010). Franz Boas and Inuktitut terminology for ice and snow: from the emergence of the field to the “Great Eskimo Vocabulary Hoax”. *SIKU: Knowing our ice*, 377-400. Dordrecht, Springer.

Lafont, C., et Peña, L. (1999). La tradición humboldtiana y el relativismo lingüístico.

Lampis, M., Strbáková, R. (2017). Acerca de la distinción entre lenguas naturales y lenguas cultivadas: ¿una falacia lingüística? *El Genio Maligno: revista de humanidades y ciencias sociales*, 21, 7.

Laygues, A. (2018). De l’influence des pratiques professionnelles sur l’enseignement des langues. *Sociolinguistica*, 32(1), 239–252. DOI : 10.1515/soci-2018-0020

Leavitt, J. (2014). Linguistic relativity: Precursors and transformations. *The Routledge handbook of language and culture*, 18-30, Routledge.

Lechevrel, N. (2008). L’écoulinguistique : Une discipline émergente. *Centre de recherche linguistique (CELITH)*, 3(1). <http://www.relq.uqam.ca/documents/Numero1Vol3-article2.pdf>.

Léo Léonard, J. (2017). Écologie (socio)linguistique : évolution, élaboration et variation. *Langage et société*, 160-161, 267-282. <https://doi.org/10.3917/lis.160.0267>

Linell, P. (2005). *The written language bias in linguistics. Its nature, origins and transformations*. Londres: Routledge, 264 p. ISBN: 970-0-203-34276-3 <https://doi.org/10.4324/9780203342763>.

Longhi, J. (2009). Les objets discursifs et le phénomène d’anticipation lexicale du discours : processus de référenciation et argumentativité dans l’activité discursive. N. Garric, J. Longhi (eds.), *L’analyse linguistique des corpus discursifs. Des théories aux pratiques, des pratiques aux théories*, 69-92, Clermont-Ferrand : Presses Universitaires Blaise-Pascal.

Longhi, J. (2018). *Du discours comme champ au corpus comme terrain. Contribution méthodologique à l’analyse sémantique du discours*. L’Harmattan.



Luzuriaga, D. (2012). *Extracción y aprovechamiento del mucilago de cacao como materia prima en la elaboración de vino*. Quito: Universidad Tecnológica Equinoccial.

Maffi, L. (2001). *On Biocultural Diversity. Linking Language, Knowledge, and the Environment*. Washington: Smithsonian Institution Press.

Mancebo-Humbert, M. ; Alves, I. M. ; et Gautier, L. (2019a). De la terminologie spontanée à une terminologie aménagée et vice-versa : parler des vins espumantes au Brésil. *XIV<sup>ème</sup> Journée Realiter*, Réseau Realiter, Paris: France.

Mancebot-Humbert, M. ; Gautier, L. ; et Le Fur, Y. (2019b). Terminologie et discours au sein d'une filière vitivinicole : Le cas des créchants de Bourgogne. *Revue des Oenologues et des Techniques Vitivinicoles et Oenologiques*, Union nationale des œnologues, France Bourgogne-Publications, 17.

Marcellesi, J.-B., B. Gardin, B. (1974). *Introduction à la socio-linguistique : La linguistique sociale*. Paris: Larousse.

Martinet, A. (1960). *Éléments de linguistique générale*. Paris: Armand Colin.

McEnery T., et Wilson A. (2001). *Corpus linguistics: An introduction*. Edinburgh University Press.

Méric, O. (2016). *Organisation discursive de la visite médiée de sites touristiques : théorisation contributionnelle et valorisation d'une praxis professionnelle*. (Doctoral dissertation, University of Burgundy).

Méric, O. (2019). Sémiotisation discursive de l'émotion : analyse textométrique d'un corpus plurilingue de discours spécialisés produits autour du chocolat. *Affects, émotions et expressivité dans les discours sociétaux*, édité par Laurent Gautier et coll., Francfort/Main: Peter Lang.

Méric, O., Carrera Rodríguez, E. M. (2019). "Aplicación de la Interpretación ambiental en el rescate de la Chakra Amazónica como recurso turístico comunitario en la Provincia del

Napo”. *Ambiente y sustentabilidad - volumen 3 Memorias III Congreso Internacional sobre Agroecología “Comunidades en armonía con la naturaleza”*. DOI: 978-9-942-932-28-0.

Méric, O. et Gautier, L. (2017). Le discours du guide-conférencier comme marqueur d’identité d’une institution touristique : l’apport d’un corpus oral authentique. *Études de linguistique appliquée*, 188, 443-465.

Montoya-Abat, B. (1996). *Alacant: la llengua interrompuda*. València, Denes editorial

Morange, S. (2009). Expert ? Vous avez dit expert ? D. Dubois (éd.) *Le sentir et le dire, Concepts et méthodes en psychologie et linguistique cognitives*, Paris: L’Harmattan : 137-156.

Mühlhäusler, P. (1996). *Linguistic Ecology: Language Change and Linguistic Imperialism in the Pacific Region*. Routledge. ISBN 9780203211281. <https://doi.org/10.4324/9780203211281>.

Mújica, C., Goldáraz, J. M. (2010). *Aprendamos Kichwa: gramática y vocabulario napeño*. Quito: Fundación Alejandro Labaka/Cicame.

Narvaja de Arnoux, E. (2021). El análisis del discurso en Latinoamérica: Objetos, perspectivas y debates. *Signos*, 54 (107), 711-735.

Nation, P. (2006). How Large a Vocabulary Is Needed for Reading and Listening? *Canadian Modern Language Review-revue Canadienne Des Langues Vivantes - CAN MOD LANG REV*, 63, 59-81.

Páez Paredes, L. (2015). *Ecuador Tierra del Cacao*. Quito : édition Trama. ISBN: 978-9978-369-575

Parizot, A., Verdier, B., et Catellani, A. (2019a). Du discours gastronomique et œnologique. *Recherches en communication*, 48, 1-5. Louvain: Centre de recherche en communication, Université Catholique de Louvain.

Parizot, A., Verdier, B. (2019b). Quand le vin se fait chair : Liminalité du discours sur le vin. *Recherches en communication*, 48, 29-46. Centre de recherche en communication, Université Catholique de Louvain.

Pazmiño Arízaga, A. L. (2015). *Estudio de factibilidad para la producción y exportación de chocolates rellenos con dulce de níspero hacia europa (bielorrusia)*. (Bachelor's thesis, Quito/UIDE/2015).

Pery-woodley, M.-P. (1995). Quels corpus pour quels traitements automatiques ? *Traitement Automatique des Langues*, 36 (1-2), 213–232.

Petit, Michel (2010). Le discours spécialisé et le spécialisé du discours : repères pour l'analyse du discours en anglais de spécialité. *E-rea*, 8, 1. DOI: 10.4000/erea.1400. <http://erea.revues.org/1400>.

Pincemin B., Heiden S., Lay M.-H.e, Leblanc J.-M., et Viprey J.-M.. (2010). Fonctionnalités textométriques : Proposition de typologie selon un point de vue utilisateur. *JADT 2010 - Statistical Analysis of Textual Data -Proceedings of 10th International Conference*. Rome: Edizioni Universitarie di Lettere Economia Diritto.

Porras, V. H., Sanchez, J. A. (1991). *Enfermedades del Cacao*. IICA Biblioteca Venezuela.

Programme de Nations Unies pour l'environnement (février 2002). *Rapport du Conseil d'administration septième session extraordinaire, assemblée générale Document officiels Cinquante-septième session supplément N° 25 (A/57/25)*. Nation Unies : New York. ISSN 0252-208X. <https://www.un.org/french/ga/57/docs/a5725f.pdf>

Pullum, G. K. (1991). *The great Eskimo vocabulary hoax and other irreverent essays on the study of language*. University of Chicago Press.

Quingaísa, E., Riveros, H. (2007). Estudio de caso: denominación de origen “cacao arriba.” *Fao-Iica*, 1–70. <http://orton.catie.ac.cr/repdoc/A7704E/A7704E.PDF>.

Rastier, F. (2004). Enjeux épistémologiques de la linguistique de corpus. *Rubrique Dits et inédits*. [http://www.revue-texto.net/Inédits/Rastier/Rastier\\_Enjeux.html](http://www.revue-texto.net/Inédits/Rastier/Rastier_Enjeux.html)

Rastier, F. (2005). Enjeux épistémologique de la linguistique de corpus. *La linguistique de corpus*, 31-45. [http://www.revue-texto.net/Inedits/Rastier/Rastier\\_Enjeux.html?iframe=true&width=100%&height=100%](http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html?iframe=true&width=100%&height=100%)

Reboul, A. (1992). *Rhétorique et stylistique de la fiction*. Nancy: Presse Universitaire de Nancy.

Recalde, M., ozas, V. V. (2009). Problemas metodológicos en la formación de corpus orales. *A Survey on Corpus-based Research/Panorama de investigaciones basadas en corpus*, 37-49.

Rodrigo-Mendizábal, I. F. (2018). Análisis del discurso en Latinoamérica: un estado de arte. Chasqui. *Revista Latinoamericana De Comunicación*, 139, 15-66.

Sharifian, F. (2017). *Cultural Linguistics*. Amsterdam et Philadelphia: John Benjamins.

Salikoko S. Mufwene, (2018). Language Evolution from an Ecological Perspective. Alwin F. Fill et Hermine Penz (eds.) *The Routledge Handbook of Ecolinguistics*, 73-88. [https://www.researchgate.net/publication/324979275\\_Language\\_evolution\\_from\\_an\\_ecological\\_perspective](https://www.researchgate.net/publication/324979275_Language_evolution_from_an_ecological_perspective)

Sapir, E. (1968). *Linguistique*. Gallimard, Paris.

Saporta G. (2006). *Probabilités, analyse des données et statistiques*. Paris: Éditions Technip, 622 p. ISBN 978-2-7108-0814-5

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. (1992). The automatic analysis of corpora. Svartvik, J. (ed.) *Directions in Corpus Linguistics Proceedings of Nobel Symposium 82*. Berlin: Mouton de Gruyter.

Skutnabb-Kangas, T., R. Phillipson (2008). A Human Rights Perspective on Language Ecology. A. M. Creese, N. Hornberger (eds.) *Encyclopedia of Language and Education*, 9, Nueva York: Springer, 3-14.

Stefanowitsch A. (2020). *Corpus linguistics: A guide to the methodology*. Berlin: Language Science Press. ISBN 978-3-96110-225-9. DOI:10.5281/zenodo.3735822. <https://langsci-press.org/catalog/book/148>.

Stork, Y. (2008). La situación lexicográfica del español hispanoamericano. Consideraciones desde la perspectiva de la ecolingüística. *El diccionario como puente entre las lenguas y culturas del mundo. Actas del II Congreso Internacional de Lexicografía Hispánica*, Alicante, 594-600.

Sukha, D. A. (2016). *Pasos hacia una norma internacional armonizada para la evaluación del sabor de cacao – una revisión de los protocolos y prácticas actuales*. International Standards for the Assessment of Cocoa Quality and Flavour. [https://www.cocoaqualitystandards.org/fileadmin/templates/CocoaQuality/Uploads/Documents-and-reports/REVIEW-Cocoa\\_Quality\\_Flavour\\_Standards\\_-\\_Darin\\_SUKHA\\_sans\\_Appendices-12May2016-ESPANOL.pdf](https://www.cocoaqualitystandards.org/fileadmin/templates/CocoaQuality/Uploads/Documents-and-reports/REVIEW-Cocoa_Quality_Flavour_Standards_-_Darin_SUKHA_sans_Appendices-12May2016-ESPANOL.pdf)

Terborg, R. (2006). La ecología de presiones' en el desplazamiento de las lenguas indígenas por el español. Presentación de un modelo. *Forum Qualitative Sozialforschung*, 7, 4, 4-06. <http://www.qualitative-research.net/index.php/fqs/article/view/167/374>

Terborg, R., García-Landa, L. (2013). The ecology of pressures: Towards a tool to analyze the complex process of language shift and maintenance. A. Bastardas-Boada, À. Massip-Bonet (eds.) *Complexity Perspective on Language, Communication and Society*. Berlin: Springer. 219-239.

Teubert, W., (1996). Comparable or parallel corpora? – *International Journal of Lexicography*, 9 (3), 238-264.

Teun A. van Dijk (1999). *Ideología. Una aproximación multidisciplinaria*. Barcelona: Gedisa.

Torruella Casañas, J., et Capsada Blanch, R. (2017). Métodos para medir la riqueza léxica de los textos. Revisión y propuesta. *Verba: Anuario Galego De Filoloxía*, 44, 347-408. <https://doi.org/10.15304/verba.44.3155>.

Vassiliadou, H., Lammert, M. (2011). Odeurs et dimension hédonique à travers le prisme des adjectifs. *Langages*, 181, 73-88. <https://doi.org/10.3917/lang.181.0073>.

Vera, J., Goya ,A. (2015) Comportamiento agronómico, calidad física y sensorial de 21 líneas híbridas de cacao (*Theobroma cacao* L.). *La técnica*, 26-37.

Villela-Petit M. (1992). Le temps dans la langue et la culture hopi. *L'Homme et la société, Anthropologie de l'espace habité*, 104, 121-136. DOI: 10.3406/homso.1992.2645. [http://www.persee.fr/doc/homso\\_0018-4306\\_1992\\_num\\_104\\_2\\_2645](http://www.persee.fr/doc/homso_0018-4306_1992_num_104_2_2645)

WHORF, Benjamin L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. MIT Press.

Wittgenstein, Ludwig ([1921] 1993). – *Logisch-Philosophische Abhandlung*. – Routledge et Kegan Paul Ltd., trad. GRANGER, Gilles Gaston. – *Tractatus logico-philosophicus*. – Paris : Editions Gallimard, collection Bibliothèque de Philosophie

Wu, Y. (2018). Ecological discourse analysis. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, 181.

Yépez, J. C. (2015). *Efecto de Altas Densidades Y Dos Sistemas de Siembra Sobre El Rendimiento Y Rentabilidad Del Cultivo de Plátano*. Quevedo: Universidad técnica Estatal de Quevedo.

Zentella, A. C. (1998). *Growing up bilingual: Puerto Rican children in New York*. Blackwell.

Zimmermann, K. (1999). *Política del lenguaje y planificación para los pueblos amerindios: ensayos de ecología lingüística*. Madrid: Iberoamericana Vervuert, 200 p. Lengua y sociedad en el mundo hispánico. ISBN : 9783865278876